

1 TCCTCCGGTC GCCCCGCCCTC GGGGCAGCTA GTGGCGCAGC CCCCCGCCCG
 51 CGGCCCTGGC CTCCCCGGCG GCGCGGCAGG GGAGGGCTTA AGCTGCCGCA
 101 GGGACCGCCG CGTGCGGGGC GAGAGGGAGC CCCGGTGGG GGTGGCGCAG
 151 CGGGCGGGGT TCGGTCCGAG CCCGGTGGGA GGCTCCCGA GCGCAGCCTG
 201 GGCCCAGCCC ACCCCCGCAGC GCGGGCATG GCAGGCACCC TGGACCTGGA
 251 CAAGGGCTGC ACGGTGGAGG AGCTGCTCCC CGGGTGCATC GAAGCCTTCG
 301 ATGACTCCGG GAAGGTGCAGG GACCCGCAGC TGGTGCAGT ATTCCATG
 351 ATGCACCCCT GGTACATCCC CTCCCTCAG CTGGCGGCCA AGCTGCTCCA
 401 CATCTACCAA CAATCCCAGA AGGACAACTC CAATTCCCTG CAGGTGAAAAA
 451 CGTGCCACCT GGTCAAGGTAC TGGATCTCCG CCTTCCCAAGC GGAGTTGAC
 501 TTGAACCCGG AGTTGGCTGA GCAGATCAAG GAGCTGAAGG CTCTGCTAGA
 551 CCAAGAAGGG AACCAGCGC ACAGCAGCCT AATCGACATA GACAGCGTCC
 601 CTACCTACAA GTGGAAGCGG CAGGTGACTC AGCGGAACCC TGTGGGACAG
 651 AAAAAGCGCA AGATGTCCCT GTTGTGAC CACCTGGAGC CCATGGAGCT
 701 GGCGGAGCAT CTCACCTACT TGGAGTATCG CTCCCTCTGC AAGATCCTGT
 751 TTCAGGACTA TCACAGTTTC GTGACTCATG GCTGCACTGT GGACAACCC
 801 GTCCTGGAGC GGTTCATCTC CCTCTTCAAC AGCGTCTCAC AGTGGGTGCA
 851 GCTCATGATC CTCAGCAAAC CCACAGCCCC GCAGCGGGCC CTGGTCATCA
 901 CACACTTGT CCACGTGGCG GAGAAGCTGC TACAGCTGCA GAACTTCAAC
 951 ACGCTGATGG CAGTGGTCAGG GGGCCTGAGC CACAGCTCCA TCTCCCGCCT
 1001 CAAGGAGACC CACAGCCACG TTAGCCCTGA GACCATCAAG CTCTGGGAGG
 1051 GTCTCACGGA ACTAGTGACG GCGACAGGCA ACTATGGCAA CTACCGGCGT
 1101 CGGCTGGCAG CCTGTGTGGG CTTCCGCTTC CCGATCCCTGG GTGTGCACCT
 1151 CAAGGACCTG GTGGCCCTGC AGCTGGCACT GCCTGACTGG CTGGACCAG
 1201 CCCGGACCCG GCTCAACGGG GCCAAGATGA ACCAGCTCTT TAGCATCCTG
 1251 GAGGAGCTGG CCATGGTGCAC CAGCCTGCGG CCACCACTAC AGGCCAACCC
 1301 CGACCTGCTG AGCCTGCTCA CGGTGTCTCT GGATCAGTAT CAGACGGAGG
 1351 ATGAGCTGTA CCAGCTGTCC CTGCAGCGGG AGCCCGCCTC CAAGTCCCTG
 1401 CCAACCAGGCC CCACGAGTTG CACCCACCA CCCCCGGCCCC CGGTACTGG
 1451 GGAGTGGACC TCGGCTGCCA AACCCAAGCT GGATCAGGCC CTCGTGGTGG
 1501 AGCACATCGA GAAGATGGTG GAGTCTGTG TCCCGAACTT TGACGTGAT
 1551 GGGGATGGCC ACATCTACA GGAAGAATTG CAGATCATCC GTGGGAACTT
 1601 CCCTTACCTC AGCGCCTTTG GGGACCTCGA CCAGAACCGAG GATGGCTGCA
 1651 TCAGCAGGGA GGAGATGGTT TCCTATTTC TGCGCTCCAG CTCTGTGTTG
 1701 GGGGGCGCA TGGGCTTCGT ACACAACTTC CAGGAGAGCA ACTCCTTGCG
 1751 CCCCCTCGCC TGCCGCCACT GCAAAGCCCT GATCCTGGC ATCTACAAGC
 1801 AGGGCCTCAA ATGCCGAGCC TGTGGAGTGA ACTGCCACAA GCAGTGCAG
 1851 GATGCCCTGT CAGTTGAGTG TCGGCCAGG GCCCAGAGTG TGAGCCTGG
 1901 GGGGTCTGCA CCCTCACCC CACCCATGCA CAGCCACCAT CACCGCGCCT
 1951 TCAGCTCTC TCTGCCCGC CCTGGCAGGC GAGGCTCCAG GCCTCCAGCA
 2001 ATCCCCCTCC CAGCAGAGAT CCGTGAGGAG GAGGTACAGA CGGTGGAGGA
 2051 TGGGGTGTGTT GACATCCACT TGTAATAGAT GCTGTGGTTG GATCAAGGAC
 2101 TCATCCCTGC CTTGGAGAAA ATACTTCAAC CAGAGCAGGG AGCCTGGGG
 2151 TGTGGGGCA GGAGGCTGGG GATGGGGTG GGATATGAGG GTGGCATGCA
 2201 GCTGAGGGCA GGGCCAGGGC TGGTGTCCCT AAGGTTGTAC AGACTCTTGT
 2251 GAATATTGT ATTTTCCAGA TGGAAATAAAA AGGCCCGTGT AATTAAAAAA
 2301 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA (SEQ ID NO:1)

FEATURES:

5'UTR: 1-227
 Start Codon: 228
 Stop Codon: 2073
 3'UTR: 2076

Homologous proteins:

Top 10 BLAST Hits

| | Score | E |
|--|-------|-------|
| CRA 1000682340958 /altid=gi 6358505 /def=gb AAF07219.1 AF043722... | 1293 | 0.0 |
| CRA 18000005086608 /altid=gi 5031623 /def=ref NP_005816.1 RAS ... | 1241 | 0.0 |
| CRA 18000005188697 /altid=gi 6755290 /def=ref NP_035372.1 RAS,... | 1202 | 0.0 |
| CRA 18000005205935 /altid=gi 7662334 /def=ref NP_056191.1 KIAA... | 618 | e-175 |
| CRA 18000005188699 /altid=gi 3928857 /def=gb AAC79700.1 (AF081... | 533 | e-150 |
| CRA 18000005152782 /altid=gi 9507035 /def=ref NP_062084.1 RAS ... | 531 | e-149 |
| CRA 18000005192860 /altid=gi 7242201 /def=ref NP_035376.1 RAS ... | 529 | e-149 |
| CRA 18000005192861 /altid=gi 4038292 /def=gb AAC97349.1 (AF106... | 526 | e-148 |
| CRA 18000005188698 /altid=gi 5032025 /def=ref NP_005730.1 RAS ... | 525 | e-148 |
| CRA 1000733831533 /altid=gi 6650545 /def=gb AAF21898.1 AF081197... | 525 | e-148 |

BLAST dbEST hits:

| | Score | E |
|---|-------|-------|
| gi 5432583 /dataset=dbest /taxon=9606 ... | 1310 | 0.0 |
| gi 9876673 /dataset=dbest /taxon=960... | 1281 | 0.0 |
| gi 11286864 /dataset=dbest /taxon=96... | 1249 | 0.0 |
| gi 11285315 /dataset=dbest /taxon=96... | 1207 | 0.0 |
| gi 5432584 /dataset=dbest /taxon=9606 ... | 733 | 0.0 |
| gi 4372300 /dataset=dbest /taxon=9606 ... | 720 | 0.0 |
| gi 12295751 /dataset=dbest /taxon=96... | 700 | 0.0 |
| gi 12288965 /dataset=dbest /taxon=96... | 599 | e-168 |
| gi 6920402 /dataset=dbest /taxon=960... | 573 | e-161 |
| gi 2005039 /dataset=dbest /taxon=9606 ... | 573 | e-161 |

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

From BLAST dbEST hits:

gi|5432583 Testis
gi|9876673 Liver-non-cancerous
gi|11286864 Brain glioblastoma
gi|11285315 Brain glioblastoma
gi|5432584 Testis
gi|4372300 B Cell Chronic lymphatic leukemia
gi|12295751 Adult marrow
gi|12288965 Adult marrow
gi|6920402 Lymph germinal center B cell
gi|2005039 Lymph

From tissue screening panels:

Leukocyte

1 MAGTLDDKG CTVELLRG C IEAFDDSGKV RDQLVRIFL MMHPWYIPSS
51 QLAAKLLHIY QQRKDNSNS LQVKTCHLVR YWISAFPAEF DLNPELAEQI
101 KELKALLDQE GNRRHSSLID IDSVPTYKWK RQVQRPVG QKKRKMSLLF
151 DHLEPMELAE HLTYLEYRSF CKILFQDYHS FVTGCTVDN PVLERFISLF
201 NSVSQWVQLM ILSKPTAPQR ALVITHFVHV AEKLLQLQNF NTLMAVVGGL
251 SHSSISRLKE THSHVSPETI KLWEGLTELV TATGNYGNYR RRLAACVGFR
301 FPILGVHLKD LVALQLALPD WLDPARTRLN GAKMKQLFSI LEELAMVTS
351 RPPVQANPDL LSLLTVSLDQ YQTEDELYQL SLQREPRSKS SPTSPSCTP
401 PPRPVVLEEW TSAAKPKLDQ ALVVEHIEKM VESVFRNFDV DGDGHISQEE
451 FQIIRGNFPY LSAFGQLDQN QDGCISREEM VSYFLRSSV LGGRMGFVHN
501 FQESNSLRPV ACRHCKALIL GIYKQGLKCR ACGVNCHKQC KDRLSVECRR
551 RAQSVSLEGS APSPSPMHSH HHRAFSFSLP RPGRRGSRPP AIPLPAEIRE
601 EEVQTVEDGV FDIHL (SEQ ID NO:2)

FEATURES:

Functional domains and key regions:

[1] PDO00004 PS00004 CAMP_PHOSPHO_SITE
cAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 3

1 113-116 RRHS
2 144-147 RKMS
3 584-587 RRGS

[2] PDO00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 7

1 27-29 SGK
2 63-65 SRK
3 126-128 TYK
4 134-136 TQR
5 269-271 TIK
6 349-351 SLR
7 506-508 SLR

[3] PDO00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 9

1 12-15 TVEE
2 63-66 SRKD
3 117-120 SLID
4 163-166 TYLE
5 339-342 SILE
6 373-376 TEDE
7 447-450 SQEE
8 476-479 SREE
9 605-608 TVED

[4] PDO00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 4

1 19-24 GCIEAF
2 249-254 GLSHSS
3 284-289 GNYGNY
4 492-497 GGRMGF

[5] PDO00009 PS00009 AMIDATION

Amidation site

582-585 PGRR

[6] PDO00018 PS00018 EF_HAND

EF-hand calcium-binding domain

Number of matches: 2

| | | |
|---|---------|----------------|
| 1 | 439-451 | DVDGDGHISQEEF |
| 2 | 468-480 | DQNQDGCGISREEM |

[7] PDO00379 PS00479 DAG_PE_BIND_DOM_1

Phorbol esters / diacylglycerol binding domain

499-548 HNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLSVEC

Membrane spanning structure and domains:

| Helix | Begin | End | Score | Certainty |
|-------|-------|-----|-------|-----------|
| 1 | 34 | 54 | 0.713 | Putative |
| 2 | 195 | 215 | 0.653 | Putative |
| 3 | 238 | 258 | 0.788 | Putative |

BLAST Alignment to Top Hit:

>CRA|18000005086608 /altid=gi|5031623 /def=ref|NP_005816.1| RAS guanyl
releasing protein 2 (calcium and DAG-regulated); calcium
and diacylglycerol-regulated guanine nucleotide exchange
factor I [Homo sapiens] /org=Homo sapiens /taxon=9606
/dataset=nraa /length=609

Length = 609

Score = 1241 bits (3176), Expect = 0.0

Identities = 608/615 (98%), Positives = 609/615 (98%)

Frame = +3

Query: 228 MAGTLDLKGCTVEELLRCGIEAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIY 407
MAGTLDLKGCTVEELLRCGIEAFDDSGKVRDPQLVR+FLMMHPWYIPSSQLAAKLLHIY
Sbjct: 1 MAGTLDLKGCTVEELLRCGIEAFDDSGKVRDPQLVRMFLMMHPWYIPSSQLAAKLLHIY 60

Query: 408 QOSRKDNNSNLSQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID 587
QOSRKDNNSNLSQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID
Sbjct: 61 QOSRKDNNSNLSQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID 120

Query: 588 IDSVPPTYWKKRQVTQRNPVGQKQRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 767
IDSVPPTYWKKRQVTQRNPVGQKQRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS
Sbjct: 121 IDSVPPTYWKKRQVTQRNPVGQKQRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 180

Query: 768 FVTHGCTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVVAEKLLQLQNF 947
FVTHGCTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVVAEKLLQLQNF
Sbjct: 181 FVTHGCTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVVAEKLLQLQNF 240

Query: 948 NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVATGNYGNYRRRLAACVGFR 1127
NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVATGNYGNYRRRLAACVGFR
Sbjct: 241 NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVATGNYGNYRRRLAACVGFR 300

Query: 1128 FPILGVLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSRPPVQANPDL 1307
FPILGVLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSRPPVQANPDL
Sbjct: 301 FPILGVLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSRPPVQANPDL 360

Query: 1308 LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPSTSCTPPPRPPVLEEWTSAAKPQLDQ 1487
LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPSTSCTPPPRPPVLEEWTSAAKPQLDQ
Sbjct: 361 LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPSTSCTPPPRPPVLEEWTSAAKPQLDQ 420

Query: 1488 ALVVEHIEKMWESVFRNFVDGDGHISQEEFQIIRGNFPYLSAFGDLQDGCISREEM 1667
ALVVEHIEKMWESVFRNFVDGDGHISQEEFQIIRGNFPYLSAFGDLQDGCISREEM
Sbjct: 421 ALVVEHIEKMWESVFRNFVDGDGHISQEEFQIIRGNFPYLSAFGDLQDGCISREEM 480

Query: 1668 VSYFLRSSSVLGGRMGFVHNQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 1847
VSYFLRSSSVLGGRMGFVHNQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC
Sbjct: 481 VSYFLRSSSVLGGRMGFVHNQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 540

Query: 1848 KDRLSVECRRAQSVSLEGSAPSPSPMHSHHRAFSFSLPRPGRRGSRPPAIPLPAEIRE 2027
KDRLSVECRRAQSVSLEGSAPSPSPMHSHHRAFSFSLPRPGRRGSRPP EIRE
Sbjct: 541 KDRLSVECRRAQSVSLEGSAPSPSPMHSHHRAFSFSLPRPGRRGSRPP-----EIRE 594

Query: 2028 EEVQTVEDGVFDIHL 2072
EEVQTVEDGVFDIHL
Sbjct: 595 EEVQTVEDGVFDIHL 609 (SEQ ID NO:4)

>CRA|1000682340958 /altid=gi|6358505 /def=gb|AAF07219.1|AF043722_1
 (AF043722) guanine exchange factor MCG7 isoform 1 [Homo sapiens] /org=Homo sapiens /taxon=9606 /dataset=nraa
 /length=671
 Length = 671
 Score = 1293 bits (3309), Expect = 0.0
 Identities = 639/670 (95%), Positives = 643/670 (95%), Gaps = 4/670 (0%)
 Frame = +3

Query: 75 GRGGVKLPQGPPRAGREGAPGGGAAG----GVRSEPGGRLPERSLGPAPAHPAPAAAGTL 242
 GRG P + +E G +G GVRSEPGGRLPERSLGPAPAHPAPAAAGTL
 Sbjct: 8 GRGTQGWPGSSEQHQEATSSAGLHSGVDELGVSEPGGRLPERSLGPAPAHPAPAAAGTL 67

Query: 243 DLDKGCTVEELLRCIEAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIYQQSRK 422
 DLDKGCTVEELLRCIEAFDDSGKVRDPQLVR+FLMMHPWYIPSSQLAAKLLHIYQQSRK
 Sbjct: 68 DLDKGCTVEELLRCIEAFDDSGKVRDPQLVRMFLMMHPWYIPSSQLAAKLLHIYQQSRK 127

Query: 423 DNSNSLOVKTCHLVRYWISAFPAEFDLNPELAEQIKEKALLDQEGRNRRSSLIDIDSV 602
 DNSNSLOVKTCHLVRYWISAFPAEFDLNPELAEQIKEKALLDQEGRNRRSSLIDIDSV
 Sbjct: 128 DNSNSLOVKTCHLVRYWISAFPAEFDLNPELAEQIKEKALLDQEGRNRRSSLIDIDSV 187

Query: 603 TYKWKRQVTQRNPVGQKCRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHG 782
 TYKWKRQVTQRNPVGQKCRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHG
 Sbjct: 188 TYKWKRQVTQRNPVGQKCRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHG 247

Query: 783 CTVDPVLERFISLNFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNFNTLMA 962
 CTVDPVLERFISLNFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNFNTLMA
 Sbjct: 248 CTVDPVLERFISLNFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNFNTLMA 307

Query: 963 VVGGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRFPILG 1142
 VVGGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRFPILG
 Sbjct: 308 VVGGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRFPILG 367

Query: 1143 VHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLLT 1322
 VHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLLT
 Sbjct: 368 VHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLLT 427

Query: 1323 VSLDQYQTEDELYQLSLQREPRSKSSPTSPSCTPPPRPPVLEEWTSAAKPKLDQALVVE 1502
 VSLDQYQTEDELYQLSLQREPRSKSSPTSPSCTPPPRPPVLEEWTSAAKPKLDQALVVE
 Sbjct: 428 VSLDQYQTEDELYQLSLQREPRSKSSPTSPSCTPPPRPPVLEEWTSAAKPKLDQALVVE 487

Query: 1503 HIEKMVESFRNFDVGDGHISQEEFQIIRGNFPYLSAFGDLDDQNQDGCISREEMVSYFL 1682
 HIEKMVESFRNFDVGDGHISQEEFQIIRGNFPYLSAFGDLDDQNQDGCISREEMVSYFL
 Sbjct: 488 HIEKMVESFRNFDVGDGHISQEEFQIIRGNFPYLSAFGDLDDQNQDGCISREEMVSYFL 547

Query: 1683 RSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRRLS 1862
 RSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRRLS
 Sbjct: 548 RSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRRLS 607

Query: 1863 VECRRRAQSVSLEGSAPS PSMHSHHHRAFSFSLPRPGRRGSRPPAIPLPAEIREEEVQT 2042
 VECRRRAQSVSLEGSAPS PSMHSHHHRAFSFSLPRPGRRGSRPP EIREEEVQT
 Sbjct: 608 VECRRRAQSVSLEGSAPS PSMHSHHHRAFSFSLPRPGRRGSRPP-----EIREEEVQT 661

Query: 2043 VEDGVFDIHL 2072
 VEDGVFDIHL
 Sbjct: 662 VEDGVFDIHL 671 (SEQ ID NO:5)

>CRA|18000005188697 /altid=gi|6755290 /def=ref|NP_035372.1| RAS,
guanyl releasing protein 2; RAP 1A protein-specific
guanine nucleotide exchange factor 1; CalDAG-GEFI [Mus
musculus] /org=Mus musculus /taxon=10090 /dataset=nraa
/length=608
Length = 608
Score = 1202 bits (3076), Expect = 0.0
Identities = 589/615 (95%), Positives = 597/615 (96%)
Frame = +3

Query: 228 MAGTLDLKGCTVEELLRGCIEAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIY 407
MA TLDLKGCTVEELLRGCIEAFDDSGKVRDPQLVR+FLMMHPWYIPSSQLA+KLLH Y
Sbjct: 1 MASTLDLKGCTVEELLRGCIEAFDDSGKVRDPQLVRMFLMMHPWYIPSSQLASKLLHFY 60

Query: 408 QQSRKDNSLNLQVKTCHLVRWISAFAEFDLNPELAEQIKELKALLDQEGRNRRHSSLID 587
QQSRKDNSLNLQVKTCHLVRW+SAFAEFDLNPELAE IKELKALLDQEGRNRRHSSLID
Sbjct: 61 QQSRKDNSLNLQVKTCHLVRWVSAAFAEFDLNPELAEPIKELKALLDQEGRNRRHSSLID 120

Query: 588 IDSVPTYKWKRQVTQRNPVGQKCRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 767
I+SVPTYKWKRQVTQRNPV QKCRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS
Sbjct: 121 IESVPTYKWKRQVTQRNPVEQKCRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 180

Query: 768 FVTHGCTVDNPVLERFISLFNSVSQWVQLMILSKPTAQRALVITHFVVAEKLQLQNF 947
FVTHGCTVDNPVLERFISLFNSVSQWVQLMILSKPTA QRALVITHFVVAEKLQLQNF
Sbjct: 181 FVTHGCTVDNPVLERFISLFNSVSQWVQLMILSKPTATQRALVITHFVVAEKLQLQNF 240

Query: 948 NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFR 1127
NTLMAVVGGLSHSSISRLKETHSHVSP+TIKLWEGLTELVTATGNY NYRRRLAACVGFR
Sbjct: 241 NTLMAVVGGLSHSSISRLKETHSHVSPDTIKLWEGLTELVTATGNSNYRRRLAACVGFR 300

Query: 1128 FPILGVHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSRPPVQANPDL 1307
FPILGVHLKDLVALQLALPDWLDP RTRLNGAKM+QLFSILEELAMVTSRPPVQANPDL
Sbjct: 301 FPILGVHLKDLVALQLALPDWLDPGRTRLNGAKMRQLFSILEELAMVTSRPPVQANPDL 360

Query: 1308 LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPCTPPRPPVLEEWTSAAKPKLDQ 1487
LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPCTPPRPPVLEEWTS AKPKLDQ
Sbjct: 361 LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPCTPPRPPVLEEWTSVAKPKLDQ 420

Query: 1488 ALVVEHIEKMVESVFRNFDVGDGHISQEEFQIIRGNFPYLSAFGDLQDQGCISREEM 1667
ALV EHIIEKMVESVFRNFDVGDGHISQEEFQIIRGNFPYLSAFGDLQDQGCISREEM
Sbjct: 421 ALVAEHIIEKMVESVFRNFDVGDGHISQEEFQIIRGNFPYLSAFGDLQDQGCISREEM 480

Query: 1668 VSYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 1847
+SYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC
Sbjct: 481 ISYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 540

Query: 1848 KDRLSVECRRAQSVSLEGSAPSPSPMHSHHRAFSFSLPRPGRRGSRPPAIPLPAEIRE 2027
KDRLSVECRRAQSVSLEGSAPSPSP H+ HHRAFSFSLPRPGRR SRPP EIRE
Sbjct: 541 KDRLSVECRRAQSVSLEGSAPSPSPHT-HHRAFSFSLPRPGRRSSRPP-----EIRE 593

Query: 2028 EEVQTVEDGVFDIHL 2072
EEVQTVEDGVFDIHL
Sbjct: 594 EEVQTVEDGVFDIHL 608 (SEQ ID NO:6)

>CRA|18000005205935 /altid=gi|7662334 /def=ref|NP_056191.1| KIAA0846
protein [Homo sapiens] /org=Homo sapiens /taxon=9606
/dataset=nraa /length=689
Length = 689
Score = 618 bits (1576), Expect = e-175
Identities = 314/597 (52%), Positives = 409/597 (67%), Gaps = 5/597 (0%)
Frame = +3

Query: 234 GTLDLKGCTVEELLRG CIEAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIYQQ 413
G+ L K T++ELL CIE FDD+G++ + L RI L+MH WY+ S++LA KLL +Y+
Sbjct: 2 GSSGLGKAATLDELLCTCIEMFDDNGELDNSYLPRI VLLMHRWYLSSTELAEKLLCMYRN 61

Query: 414 SRKDNNSLQVKTCHLVRWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLIDID 593
+ ++ N ++K C+ +RYWI FPAEF+L+ L +E + + Q G +H SLIDI
Sbjct: 62 ATGESCNEFRLKICYFMRYWILKFPAEFNLDLGLIRMTTEFREVASQLGYEKHVSЛИDIS 121

Query: 594 SVPTYKWKRQVTQRNPVGQKKRKM SLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFV 773
S+P+Y W R+VTQR V KK K LLFDHLEP+ELAEHLT+LE++SF +I F DY S+V
Sbjct: 122 SIPS YDWMMRVTQRKKVS-KKGKACLLFDHLEPIELAEHLTFLEHKSFRRISFTDYQS YV 180

Query: 774 THGCTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLQLQNFNT 953
HGC +NP LER I+LFN +S+WVQLM+LSKPT QRA VIT F++VA+KLLQL+NFNT
Sbjct: 181 IHGCLENNPTLERSIALFNGISKWVQLMVL SKPTPQQR AEVITKFINVAKLLQLKNFNT 240

Query: 954 LMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRFP 1133
LMAVVGGLSHSSISRLKETHSH+S E K W +TELV++ GNY NYR+ A C GF+ P
Sbjct: 241 LMAVVGGLSHSSISRLKETHSHLSSEVTKNWNEMTELVSSNGNYCNYRKAFADCDGFKIP 300

Query: 1134 ILGVHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSRPPVQANPDLLS 1313
ILGVHLKDL+A+ + PDW + ++N KM QL L EL + + ++ N DL++
Sbjct: 301 ILGVHLKDLIAVHVIFPDWTE--ENKVNIVKMHQLS VTLSELVSLQNA SHHLEPNMDLIN 358

Query: 1314 LLTVSLDQYQTEDELYQLSLQREPRSKSSPSTSPTCTPPRPPVLEEWTSAAKP KLDQAL 1493
LLT+SLD Y TED++Y+LSL EPR+ SPTSPT+ P +P V EW PK D +
Sbjct: 359 LLTLSLDLYHTEDDIYKLSL VIEPRNSKS PSTSPTT---PNKPVVPLEALGVMPKPDPTV 415

Query: 1494 VVEHIEKMVESVFRNF DVG DGHISQEEFQIIRGNFPYLSAFGDLDQNQDG CISREEMVS 1673
+ +HI K+VESVFRN+D D DG+ISQE+F+ I NFP+L +F LD++QDG IS++EM++
Sbjct: 416 INKHIRKLVESVFRNYDHHDHG YISQEDFESIAANFPFLDSFCVLDKDQDGLISKDEMMA 475

Query: 1674 YFLRSSSVLGGRM--GFVHN FQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 1847
YFLR+ S L +M GF+HNFQE L+P C HC + GI KQG KC+ CG NCHKQC
Sbjct: 476 YFLRAKSQLHCKMGP GFIHN FQEMTYLKPTFCEHCAGFLWGIKQGYKCKDCGANCHKQC 535

Query: 1848 KDRLSVECRRAQS VSL---EGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPPA IPL 2009
KD L + CRR A++ SL GS P + F F G R AI L
Sbjct: 536 KDLLV LACRRFARAPSLSSGHGS LPGSPSLPPAQDEV FEPGVTA GHRD L D SRAITL 592 (SEQ
ID NO:7)

>CRA|18000005188699 /altid=gi|3928857 /def=gb|AAC79700.1| (AF081196)
calcium and DAG-regulated guanine nucleotide exchange
factor II [Rattus norvegicus] /org=Rattus norvegicus
/taxon=10116 /dataset=nraa /length=795
Length = 795
Score = 533 bits (1358), Expect = e-150
Identities = 267/590 (45%), Positives = 390/590 (65%), Gaps = 12/590 (2%)
Frame = +3

Query: 156 GVRSEPGGRLPERSLGPAHPAPAAMAGTLD-----LDKGCTVEELLRG CIEAFDDS 308
G R+ P GRL +S PA ++A L KG ++++L+ CI++FD
Sbjct: 17 GSRAGPKGRLEAKSTNSPLPAQPSLAQITQFRMMVSLGHLAKGASLDDLIDSCIQSFDAD 76

Query: 309 GKV-RDPQLVRIFLMMHPWYIPSSQLAAKLLHIYQQSRKDNSNSLQVKTCHLVRYWISAF 485
 G + R QL+++ L MH I S++L KL+++Y+ + + NS + +K C+ VRYWI+ F
 Sbjct: 77 GNLCRSNQLLQVMLTMHRIIISAAELLQKLMNLYKDALEKNSPGICLKICYFVRYWITEF 136

 Query: 486 PAEFDLNPELAEQIKEKALLDQEGNRRHSSLIDIDSVPVTYKWKRQVTQRNPVG-QKKRK 662
 F ++ L ++E + L+ G H LID + + W R++TQR KKRK
 Sbjct: 137 WIMFKMDASLTSTMEEFQDLVKANGEESHCHLIDTTQINSRDWSRKLTQRIKSNTSKRK 196

 Query: 663 MSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHGCTVDNPVLERFISLFNSVSQ 842
 +SLLFDHLEP EL+EHLTYLE+SF +I F DY +++ + C +NP +ER I+L N +SQ
 Sbjct: 197 VSLLFDHLEPEELSEHLTYLEFKSFRRISFSDYQNYLVNSCVKENPTMERSIALCNGISQ 256

 Query: 843 WVQLMILSKPTAPQRALVITHFVVAEKLLQLQNFTLMAVVGGLSHSSISRLKETHSHV 1022
 WVQLM+LS+PT RA V F+HVA+KL QLQNFTLMAV+GGL HSSISRLKET SHV
 Sbjct: 257 WVQLMVLSPRTPQLRAEVFIKFIHVAQKLHQLQNFTLMAVIGGLCHSSISRLKETSSHV 316

 Query: 1023 SPETIKLWEGLTTELVTATGNIGNYRRRLAACVGFRFPILGVHLKDLVALQLALPDWLDPA 1202
 E K+ +TEL+++ NY NYRR C F+ PILGVHLKDL++L A+PD+L+
 Sbjct: 317 PHEINKVLGEMTELLSSCRNYDNYRRAYGECTHFKIPILGVLKDLISLYEAMPDYLEDG 376

 Query: 1203 RTRLNAGAKMKQLFSILEELAMVTSRPPVQANPDLLSLLTVSLDQYQTEDELYQLSLQRE 1382
 ++N K+ L++ + EL + + PP+ AN DL+ LLT+SLD Y TEDE+Y+LS RE
 Sbjct: 377 --KVNVQKLLALYNHINELVQLQDVAPPDANKDLVHLLTLSLDLYYTEDIYELSYARE 434

 Query: 1383 PRSKSSPTSPSTSCTPPPRPPVLEEWTSAAKPKLDQALVVEHIEKMVESVFRNFVDGDGH 1562
 PR+ +P P +PPV+ +W S PK D + +H++MV+SVF+N+D+D DG+
 Sbjct: 435 PRNHRAPP----LTPSKPPVVWDWASGVSPKDPKTIKHVQRMVDSVFKNYDLDQDGY 489

 Query: 1563 ISQEEFQIIRGNFPYLSAFGDLQDQNQDGCIISREEMVSYFLRSSVLLGG-RMGFVHNQES 1739
 ISQEEF+ I +FP+ +F +D++++G ISR+E+ +YF+R+SS+ +GF HNFQE+
 Sbjct: 490 ISQEEFEKIAASFP--SFCVMDKDREGLISRDEITAYFMRASSIYSKLGFLPHNFQET 547

 Query: 1740 NSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLSVECRRRAQS 1889
 L+P C +C + G+ KQG +C+ CG+NCHKQCKD + EC++R++S
 Sbjct: 548 TYLKPTFCNDNCAGFLWGVIKQGYRCKDCGMNCHKQCKDLVVFECKRSKS 597 (SEQ ID
 NO:8)

Hmmer search results (Pfam) :

| Model | Description | Score | E-value | N |
|---------|--|-------|---------|---|
| PF00617 | RasGEF domain | 123.5 | 4e-33 | 1 |
| PF00130 | Phorbol esters/diacylglycerol binding domain | 59.5 | 3.6e-14 | 1 |
| PF00036 | EF hand | 21.8 | 0.00027 | 2 |
| PF01237 | Oxysterol-binding protein | 3.5 | 4.2 | 1 |

Parsed for domains:

| Model | Domain | seq-f | seq-t | hmm-f | hmm-t | score | E-value |
|---------|--------|-------|-------|-------|----------|-------|---------|
| PF01237 | 1/1 | 249 | 272 | .. | 1 24 [. | 3.5 | 4.2 |
| PF00617 | 1/1 | 148 | 336 | .. | 1 227 [] | 123.5 | 4e-33 |
| PF00036 | 1/2 | 430 | 458 | .. | 1 29 [] | 17.4 | 0.0047 |
| PF00036 | 2/2 | 463 | 482 | .. | 5 24 .. | 6.7 | 4.9 |
| PF00130 | 1/1 | 499 | 548 | .. | 1 51 [] | 59.5 | 3.6e-14 |

1 ACAGAAAGGT CCTGTTCTA AGTCTTACAT TACCAAGACT GAGGTGCGGG
51 GGCGGTCTG GATCCCCCGC CCCAAGGCTG GGAGGGGCAC GCCTCGGAAG
101 GGAGGTTTGG GGTCGGTGGT TTCACAGTGA GTGTGTCTGA AGCCAAATGG
151 TCGGAAACCG TTACCCGCTC TCCTAGGCC GGCTAGTGGG GACCCCAACC
201 GCCTGCGGCT GCCCCCTCCA AGTCCCTCCC TGTTGGCCAG GCATCCAGGT
251 CTCCAGTCTC CGAGCTGCGG AGAACCCACC GCCACATGCG GCTGCCCCTT
301 TCCATTGAC CCTGTGGGA GCCAGGCTTC CGGGGCCCG TTCCCTCTGT
351 GTGAACTGGG CCCCCCGCCC CCATTCCCAG ACATCAAGGC CGCGTCTCCA
401 GATAGCCACG ATTTCATTC TCGCTCCCA CAGGTCCCTC TCCCCAAAAT
451 ATTCCCATCT TGTCTTAGCC CATCCCCAG ACTATCTCAA GGACCAAGCTG
501 TCCCCACGCG CCCGACCTCC ACTAGGGCTG TGCCACCGC TGCTCGCAGG
551 AAGACGCCCG GTCCCGGGCC GGGTTAGCCC CATGGGAACG GTTGTCTCG
601 AAAACAGGAA CCCGAGCTGG GGGCTGGCG GGGGCCCGT TCCCCACCGC
651 AGTCCGCTTC CTGCCCCCTC CGGCTTCCCTC CGCCCCGACAC CCAGGCAGGG
701 CGGGGGCAC TGGGGCGTCC GCGGTTGGG GAGGGGCTCT TCCTTTCGGT
751 CCCCCCTCCC CGGTCCCAGG CGGCAGGGCC TCCGGTCGCC CGCTCGGGG
801 CAGCTAGTGG CGCAGCCCCC CGCCCGCGGC CCTGGCCTCC CGGGCGGC
851 GGCAGGGGAG GGGTTAACGCT GCCGCAGGGG CCGCCGCGTG CGGGCGAGA
901 GGGAGCCCCC GGTGGGGGTG GCGCAGCCGG CGGTGCGGAG CTCCGCGCAG
951 GGGCGGAGGG GGGAGGGGGC AGCCTGGCGC GGGGGCGGGG GCGGGGGC
1001 GGGGAGCGGG GCGCGGGCGT GGAGAGCGGG CGGGAGCCGC AGCCGAGCG
1051 AGGCCGGCGG CGGGGAGCGC ACGGAGGTGG GTCGGCCAG GCGGTGCGG
1101 GCTCCTGCG GCAGGTCCCA AGAGTGAATG GGCAGCGCG GCGGGGGCGC
1151 CAGGCGAAGG AGGGCGCGG CCCCAGCGAC TCCCCCCCCCG CCCAGGGCG
1201 CGCGGGCGGG CTGGGGCGG CGAGCGGGTG GGGAGTCTGC GGCCCGGGTC
1251 TGGGAGAGGG GGCAGCGGCC ACAGAGAGCTA AGGCGCGCTG GATCCCCGA
1301 GGGCGGAGGA CCTCCACGGT GCACCCAGCT TTTCCAGCC ACCTTCCAGC
1351 GGGGCCCTCC CCCCGTACCC CCCATTGGC AGATGAGAAA ATTGAGGCTC
1401 CCAGAGGCCA AGTGAATCTC AAGGTCACAC GAGGAAGCGGG TAGAGCCAGG
1451 CGGGGACGGC TCTGGGTGGC TCTTAGGAAA AGTCCGCTG AGAACTCCGT
1501 ACAGGGAGCTC CCTGTCTC CAGCCTGGG GAGTGAATG ATGTAGGGCC
1551 GGGGTACCTT TCGTGGGG AAGGCTCTGC CAAAATCTGG GAGTGAAGGG
1601 AGTCAGGGAG CTGGGGCGG AGGGCGGGCC CTGCACCGCA AATGGGAGGG
1651 GGGCAGCGGA ATGGCGTGC GCACCCATGG GGGTGTGTGC ATGTGTGTGG
1701 GAGTGTACAT GCGTGGAGAG GCACTGCCTT GCGTGTGTGC ACACGTGTGA
1751 GGATGTCAGC GCCTGTGTGG CGCGGGACT CAAGGCTGGC CTGGCTCAAG
1801 TGAACAGCAC GTCCAGGGAG CGACCTCGTC CGCGGGTTTG CATTCTGGG
1851 TGGACGAGCT GGGTATGTG GCCTGAGGGT TTCTTCGTGC AGGTGTGCAC
1901 AGGGTGTGGG TGCCATTGTG TGTGAGAGAC GGAGGATGGG GAGGCCGGTG
1951 CCTGTGGCCC GGTGCGTGTG AGTGGGACG CCTGCACCTC CACTTAGTC
2001 CCCGGCTCC GACGACTAAC TTGGGTGTGG AGTGTGGGCC CCTGCCAGGG
2051 TCGGTATGAC CCCGCCAGTG ACCGGAGTTG CTAATGGGT CATGCACCCA
2101 CCGGCCACCC TTGGCGCGAG CGCCCCCTC TGGACACCCCT GCTCCGTGCG
2151 CGCTCACAGT TCGCCTGTG GGGGCCGGGG CCAGGGTCAG GAGCCGGGA
2201 TAGGGAGGAA GAGGGCCTGT GGACAAGCTG AGCCGGGACCC CTTGGGACCT
2251 TTGCGGAGGT GGCCTGGGAG CGCTCAGTTC CCAGGCTGAG GCTTCCCCT
2301 GACGCCCTCT GGCGCAGCG GGCTCCCCC GCCCCAGGAA TGTTCCCTCTC
2351 CCATCCAGTC CGCCTCCCCCT AGGGCAGGCC CCCTGGGGC TGCCGAGCC
2401 CCCGCTCGCC TTCCCTGGGCT CCCGGGAGGG GGCAGGGCGA GCAGGACGCC
2451 TGGTTCTCT CCCCCCACCT CCCATACCAAG GGAGAAATTCTC CTCCGAGGTC
2501 CCCTCAGGCT CTGGGTTCCC AAAATAACCC TGCGGGGGAA GGGAGGCTGT
2551 GGAGGGAGGG AAGCGGGAGG GGCGCAGAGC CGAGCTGCGG GGTGCTGCAG
2601 GTGCGCTCTGG GGAGAGGGCG CGAGGAGAAAG GGCCTCGC GGGGGCTGGG
2651 CGCCAGGCCAG TCCGGGATC TTGGTTCGTC CCCATCCTCG TGAAGCCCT
2701 CGGCCCTCCC GCGACTCCGA GGGTGGGCCGG GAAGCCTCTC TGCGGGTCCG
2751 TTTCCCAACT GGGGGTTTC ACCATCCCCG GCCAGACCGT TTAACCCCCG
2801 GAGTGGCCGC GGGGGACAAAC TCCGCCCCCTG TCCAGCAGGG GGCCTGCC
2851 CCCCGCCCCG TTTCTGCCCC CGGGGCCGCT CCCCGCCCCG CGACTCCGCA
2901 GACTCCCGCT CTGCCCTCTCC CGGGACAGGG GTTCGGTCCG AGCCCGGTGG
2951 GAGGCTCCCG GAGCGCAGCC TGGGCCCAGC CCACCCCGCG CGGGCGGCCA
3001 TGGCAGGCAC CCTGGACCTG GACAAGGGCT GCACGGTGGA GGAGCTGCTC
3051 CGCGGGTGCA TCGAAGCCTT CGGTGAGTGG CTCGGGAGGG CACACGGAGC
3101 CTGAGCCTAG CCCCGAGTCT GAGCCGGGT CCCTGCCTCC CAGGCACAGT

3151 CCAGGGCACA GCCCTGACCC GGACCCACCC TGCTCCGCAG CGTGCAGTCT
3201 CTTAACGAA AGCCTCCTCC GCAACGCAGG GCAGAGAGAT GCACGCCCTT
3251 CAGACAGATG AGGTTTCCCT TCTCTAGCCT TCCCCAGCGG CGGCGAAGGG
3301 AGGGCCGGT CCCGGACTCT GACACTTGAG GGGCATTATC TGTCTCCCGG
3351 GGAATCCGGA GGAACTCGCT ATCTCCGGCC TGGGAGCTGT TTCCGGCTAA
3401 TGGGGGGCGG CTTATCTGGT GAAGGGGTGC CCCCTCCCCC CAAGGGCTCA
3451 GGAATGACC TCTGGATTCT TGACCCCGGG GAACCCAGGC TCCTTCCGCC
3501 CCAGCTGGTT CCCCTCCGGA CGATGGCGG CTCGGGCGCT CCCCTCCCTC
3551 AGTCCTCAGG GCGTGCCTAT CTCTCGGCCA CCACACCTTT CCTCTCTAAT
3601 TTGCTCTCTG CTCTCGGAGT CTCGGGCAAG CAGGAGGTGG CGGGGGTGA
3651 GCGTGCACCC GAAGGACCGA TACCTGGCGG GTTGGGGGT GAGGATGAGG
3701 CATGGTAGCT CGGGACCCAG CTCAGCCACC TGTCTTGAC CCTTCGGAGT
3751 CAGATGACTC CGGGAAAGGTG CGGGACCCGC AGCTGGTGC CATGTTCTC
3801 ATGATGCACC CTCGGTACAT CCCCTCTCT CAGCTGGCGG CCAAGCTGCT
3851 CCACATATAT CTTTCGCCGG CCTTGCCAAG GCCCCCCGCCG TCGGAGCCCA
3901 TGCGCAGCCC CTCTGCCAG CCCAGGTGCA GAATGAGCCT CGCTCTAAG
3951 TATAGGCCAC TCCTTATCCC AGAGCTCAGG CGTCGTCCTA GCCTCCAAT
4001 AGGGCCTAGG CTCTGCCCTC TCCTTGCTCC TAGCGACTCG GTCTGTCCC
4051 CAGGCTCTGT CCCCAGCCGA GGCCTTGCCC TCCTTCTCCC TAGAGTCTAG
4101 GGCCTGCCCT TGCTTCAGGG TTGGGTGCGC CCCGTGCATC TCTCTCTCCC
4151 AGAGCCCAAGG CTTTGCTTC AGCCTCCCTC AGCACCTAGT CCTCCACCC
4201 CACCTCCAAC CCCTCCCAGA GCTCAAGCCT CACCCCCAGC ATCTCCGCAG
4251 AGCGCAAGCC CCATCCCTAG AACGTGTCCTC CTAGAACAG GCCCCGCC
4301 CAGCCTCCCT CCACGCAGGC CTCCCTTCT AGAGTTAACG GGCCTCTTA
4351 ACCCTCTCCT TCACCTTACCA ACAATCCCGG AAGGACAAC CCAATTCCCT
4401 GCAGGTGAAA ACGTGCCACC TGGTCAGGTG AGTCTTTCCC CTGGGGCTCT
4451 AGCCCTCCC CTTTCTCCCT TCTCTCTGGC TTCAAGGCTGG CCTGGAGGAG
4501 GGGGCAGGGC GCTGTTCTG GGAGTGGGT TGAACCCCTGG CTTGTCGGGG
4551 TGGGCAGTGC TGCCACAGGC TCACCCCTTC CTGGGTCTGG GCCTTAATT
4601 TCTTTCTGC CGAGTGCAGG TGTTGCTCTC AAGGGTCTAA TGTACACTG
4651 GAGTGGCGAA GAAAGAGCT GGAACCATAG TTGAGGGTC TTTTGCTTA
4701 GGTGACTATA ATCTCAAATA GTCCTTGCA GCCTGCTGGG TGATGGTGGG
4751 GGAAGGGCTA TCTGGGTGA CTCCCCGCTC CTCCAGGTAC TGGATCTCCG
4801 CCTTCCCAGC GGAGTTGAC TTGAACCCGG AGTTGGCTGA GCAGATCAAG
4851 GAGCTGAAGG CTCTGCTAGA CCAAGAAGGG AACCGACGGC ACAGCAGCCT
4901 AATCGACATA GACAGCGTGT GCGTGGGGGG AGCACAGAGG GCTGGGGGG
4951 CACTCAGTAT CCTATACCAT CTGTGCTTAA TAAATGTCTG TTGAACTGAA
5001 TGAGTGAGGG TCATGTTGCT CTCTCGCTTA AAAACCTTCC ATGGCTCCCT
5051 ATTGCTCTCA ACATGCCTCC TCTGGCAGC TTGGCGTTC TGCTCATCT
5101 TCCACTGCCA CCACCCATCC CACACACCTC CTCCCTGTAGC TGCGCTGGGT
5151 CGGCTCCCCG TCGGCTGAGC TCTCGAGTCC TTTCTCATCA TGGTGCTCTG
5201 CTCATATCAT CCCCCCTGCT GCCTCCCTCG TGTACCAAG ACTCAGTCA
5251 GGCATGAAGT CTCCGTGGC TCTGAGGGTT CGGGGCTCTT CGGGGGTAGA
5301 ATTGTCGTT CCCACCTCTG TTTCCATGG CACTTGTAC AGACTCCTGT
5351 ACAAAAGACCT CTGTACATGT GTCACCGCTGT TTGATGATCA TGTGTTCTG
5401 TGTCTGTCTC CCTCAGTAGA CTGTGAGCTC CTGAGGGCA GGAACCGTGT
5451 CTTACTCATC TCTGTATTC CAGCGCCTAG CACAGTGCCT GGCACAGAGT
5501 ACGTTGTTCA TAAATGTGTG TTGAGTGCAT GACGGGGTGG GGGGAGATGA
5551 GGAGGAGTTG CTGGGACTGG GAACATTCGT GCCTAGGACA GTGCCTCGCA
5601 TTATGTAGGT TCTCAGTAAG CGTGAATGGT GTGTCTGTGT GAGTGGGGGG
5651 CCACGAGGCA TGCGCATGTC CAGCAAAGGG CTCACTACCC CTGCCCCCCC
5701 AGCCCTACCT ACAAGTGGAA GCGGCAGGTG ACTCAGCGGA ACCCTGTGGG
5751 ACAGAAAAAG CGCAAGATGT CCCTGTTGTT TGACCCACCTG GAGCCCATGG
5801 AGCTGGCGGA GCATCTCACC TACTTGGAGT ATCGCTCCTT CTGCAAGATC
5851 CTGGTGCAGC CCGAGGGCTG GGGGGTCAGG GGTCAAATGT GGGCTGGAAG
5901 AGAGTTCTAG GAGGGGCAGG GTCCCTGGCG TAGGCTGGGT CACAGGGTGC
5951 ATCAGGGGTT TCAGTGTAAAC CACTGAAGGT CAGTGGAGG GTGAGGAGTG
6001 GCTATCAGTG AGGGGAGAGG CCGGCAAGGT GCTGAGGCCA CTCCCTCATGC
6051 CCCAGTTTC AGGACTATCA CAGTTCTGT ACTCATGGCT GCACTGTGGA
6101 CAACCCCCGTC CTGGAGCGGT TCATCTCCCT CTCAACACGC GTCTCACAGT
6151 GGGTGCAGCT CATGATCCTC AGCAAACCCA CAGCCCCGCA GCGGGCCCTG
6201 GTCATCACAC ACTTTGCTCA CGTGGGGAG GTGCCTGCCCT CTCCTCCCC
6251 GTGTCTCCCA ACCACCCAC ATGCCAGTCA GGCCAACCCT TCCCTCCCC

6301 TAACCCACTG CCTTCTCTCT AGATAAGCTG GGCCAAATTC TGGGCCACT
6351 CAGTGAATCC CTGCTCTC GTCCCCATT GCCTTCAGA AGCTGCTACA
6401 GCTGCAGAAC TTCAACACGC TGATGGCAGT GGTCGGGGC CTGAGCCACA
6451 GCTCCATCTC CGCCTCAAG GAGACCCACA GCCACGTTAG CCCTGAGACC
6501 ATCAAGGTGC CTGGGACTGG GGAGGGCCG GTGTTCCCA GGTCTGTCTT
6551 CACTGGTCC TCCCAGCAGC ACTGGGGCT GGGCACAGCT GTCCCTCATTT
6601 GATAGATATG GAAATGGAGG CTCAGAGGGG TTAAGTGCTT TTCTCAGTTT
6651 GCACAATGGC AACAGCAGAG TGGGGCTCA CAGGTCGTCA GGGACCCAA
6701 AGCTAGTACT TTTTTTTTT TTTTTAAGAC AGGGTCTCTC TCTCTGTTGT
6751 CCAGACTGGA GTTCAGTGGT GCAGTCACAA GCTCACTGCA GCCTGAATT
6801 CCTGAGCTCA ATCGATCCTC CCACCTCAGC CTCCTGAGTA GCTGGGACTA
6851 CAGGTGTACG CCACCATGCC TAATTGTTGT ATTGTTATTAA ATTTTTTTTT
6901 TTTTTTTTA GAGATGGGGT TTTGCCATGT TGCCCAGACT GGTCTTGAAAC
6951 TCCTGGGCTC AAGTGATCCG CCTGCCCTGG CCTCCCAAAG TGCTGAGATT
7001 ATGGCTTGAG CCATTGTCGC TTGCCACTTG TAGTTCTTC TTTTCTTCT
7051 CCTTCATTTT TTATTATTT TGAAGTATT TGAAGTATTG AGTAACATAC
7101 ATATAGAAAA GTATATAAAA ACATATGAGA CTGGGCGTAG TAGCTCACAC
7151 CTGTAATCCC AGCACTTTGG GAGGCTGAGG TGGGAGATC ACGTGACATC
7201 AGGAGTTGAG GACCAGCCTG GCCAACAAAG TGAAACCCA TCTCTACTAA
7251 AATACAAAAA TTAGCCAGGC ATGGTGGCAC GCACCTGGAA TCCAAGCTAC
7301 TTGGGAGGCT GAGGCAGGAG GAGAATTACT TGAACCTCAGG AGGCGGAGGT
7351 TGCAGTGTGAGC CAAGATTGTG CCACCTCACT CCAGCCTGGG CGACAGAGTG
7401 AGACTCCATC TAAAAAAAGA GAAAAGTATA TAAAAACATA TGAATAGTTT
7451 AAAGAAAAAT TGTAAGAGAA ACACTGTGTA ACTACTGCC GGGTTGGGAA
7501 ATAGAACCTT GCCAGGCCCG CAAGGCCCA GCACCTTAA GCATAACTCC
7551 CTCCCCACGA CTTTGCAAT GATGATCTT CTTTCTTTA TAGCTTCACC
7601 ATGTAGGTAT GCGGTCCAAA ACAATGTGGG GCTTTTGTG GTCTGTTTG
7651 AACTTCTAT GAATGGAATG TTGTTGTTGT TATTTTATGT CTTGTTTTT
7701 TCATTCACA TGGTCTGAG AGTCTTTCA TTCTGTCTG TGGAGCAATT
7751 GTTTTTCAT TTTCATGCACT ATATAAATT TTATGTACG TCTACCCAA
7801 TTCATTATT TATTTATTT TTTGAGATGG AGTCTGTCTC TGTCATCCAG
7851 GCTGGAGTGC GGTGGCAGA TCTCATCACT GCAACTTCCG TCTCCTGGGT
7901 TTACGTGATT CTCGTGCTC AGCCTCTGA GTAGCTGGG TTATGGGCTC
7951 GTACCAACCAC GTCTGGCTAA TTTTTGTTAG AGACAGGCTT TCACCATGTT
8001 GCCGAGGCTG GTCTTGAACCT CCTGAGCTCA GGCACATCCAC CCGCTTAGC
8051 CTCCAAAGT GCTGGGATTA CAGGTGTGAG CCACCTGCC CAGCCTACCC
8101 CAATTATGT ATTGATTCTA TTGTTGAATG TTGGGGTTTT TCCTTTCTT
8151 TTCTTCTTT CTTTTTCTTT CTTTTTTCTT TTTTTTTGGA GAGGGAGTCT
8201 TGCTCTGTGCG CCAGGCTGGA GTGCAGTGCAC GCTAATTGGG CTCACTGCAT
8251 CACTGCACCC TCTGCCTCCC GGGTTCAAGC GATTCTCCTG CCTCAGCCTC
8301 CTGAGTAGCT GGGACTACAG GCATGCACCA CCACACCGG CTAATTGGG
8351 TTTTTTTA GTAGAGATGA GGTTCCACC ATGTTGGCCA AGATGGTCTC
8401 CATCTCTGA CCTCATGATC CATCTGCCAT GGCCCTCCAA AGTGTGAGA
8451 TTACAAGTGT GAGCCACCA GCCCCAGCTGG TTTTCCAGT TTTTGCTGTT
8501 TGGACGGGT GGCTGAGTAT GTTCTTCCAG GTCATTGTC TGTGCTGCCT
8551 TGCCTCCCTG AGCCTCTGTT TCTCTGTAA AATGTTGATG ATTCCCTGCA
8601 TCCAGGCCCTG GTTTAGAGGT GTGGTGCTT TGCGAGTGCAG TATTGCCCTG
8651 AATTCTATGGC ATGAAATTCA ATCCCCAGGG GCTGAGAGAG CCAGTCGTGG
8701 GGGACAGTAA GGGAGGTTT TACTCTTCA CCTGCTCCCTG ACCCTGACTC
8751 CTCCTCACCC CCTCTACAT TTCCAGGGCT GAGGTAGGGGA GGATAGTTGT
8801 GGGGGTATGA CTCTCTGTGCT CTTTGTCCCC AGCTCTGGGA GGGTCTCACG
8851 GAACTAGTGA CGGGCAGAGG CAACTATGCC AACTACCGGC GTGGCTGGC
8901 AGCCTGTGTG GGCTTCCCGT TCCCGATCCT GGGTGTGCAC CTCAAGGACC
8951 TGGTGGCCCT GCAGCTGGCA CTGCCTGACT GGCTGGACCC AGCCCGGACC
9001 CGGCTCAACG GGGCCAAGAT GAAGCAGCTC TTTAGCATCC TGGAGGAGCT
9051 GGCCATGGTG ACCAGCCTGC GGCCACCAAGT ACAGGCCAAC CCCGACCTGC
9101 TGAGCCTGCT CACGGTGAGG AGCAGGGGGC AGGGAGGTGG GGAGCTGGG
9151 ACCAGGGTT GACAGTTTCC CCAGGTCTG GCTGTGGGCG TGGCCTGGGG
9201 CTCGGGTTG TGGCCAAGAA ACTGAGATCT AGCGTGGGCT CTGGGGTTG
9251 GAGTGGATGC TGAGAAGGGG TCCAGGCTCT GTGGGGGCT GTGGACTGAG
9301 GTCTGATCTC CAGGCTGGTA TGTGGACTGT GGGCAGTTG AACTGGGCCT
9351 GGGTCCCGGG TTGAGTTCTG GCAATGGGCT GTGTTCTAGG GCTGGGCCAA
9401 GCTCTGCATT CTGTGGCAG GGGTGGTTTC TAAGCATGGC CCTGGGCTCG

9451 GAGTGAAGTT CTGGGCTTGG CTTTACACTT GGTCTTGGGG TCTAGGGTGG
9501 GAGTGGGTT CTGGTTTAGA TCCAGACAAG GTTCTAGACA TTGGGCTGG
9551 GCTTAAGTGT TAAGGTTGG AGTGGATTCT TAGCTGCTTC TGGGCTCTGG
9601 AGGGGATCAG GTTGAAATC AGAGCTTCTG GCTGGGTTCC GACCTGGCTT
9651 CTTCCCTGAC ATCTTGGCAA TATGTTGTGT TCAAGGTTTG GGGCCATGCT
9701 GTGGTTGAT CTGTGCGCTG GGATGACATG GGGGTTGCTG TGCTGTGTT
9751 TAAGCCAGGC TTTGTCTGA GTCTAGCTTC TGACCCGAGC TCTGGCTGAG
9801 CTGTCGCCTC TAGGTCGACCC TTTGGCCCTG GGCTCTGTGG CCGTGGGAG
9851 GGGCCAGTGG GGGTGATCAG ATCTGTGTGT CCCAGGTGTC TCTGGATCAG
9901 TATCAGACGG AGGATGAGCT GTACCAAGCTG TCCCTGCAGC GGGAGCCGCG
9951 CTCCAAGTCC TCGGTGAGGG GGTACTCCCT CCTCTCCACT CTGCCCTTCC
10001 CTCCTGAGAA TCCCAGGATG TGAGGATGGG AAGAGCTCTT AGCAGCCACC
10051 TCACCCATCC ATCTTGTAGG ACAGAGGCAT CCTGGGGGTA GGGCAGTAGT
10101 GTTGGGCAGA CTTCCCTCTC CCAGGGATTG CCCCCTCTGT TCCCCGGGGC
10151 TCTGGGCTCC CCCTGCCTCT GGCCCTAGCT CAGGCCCGAC CATTTCATA
10201 GCCAACCCAGC CCCACGAGTT GCACCCACC ACCCCGGCCC CCGGTACTGG
10251 AGGAGTGGAC CTCGGCTGCC AAACCAAGC TGGATCAGGC CCTCGTGGTG
10301 GAGCACATCG AGAAGATGGT GGAGGTGAGC TCCGTGCGGAG CCTGAGCAGT
10351 GTGTGGGAG AGGCCAGTTT GCCGGAGCAC TGCCCTGGAA GCCAGCACGA
10401 GTGTCTGTGTT CAAGACCCAG CACTCAGCCC CTAGGAGTCA CAGGGCTGG
10451 CAGGCCAGCT GCACGGGGCT GAAGTGCCTCC TGGGTAGGGT GGGGGTGGAG
10501 GTATGGAACG GGGGTGGTGT CAGAGACCTC TCTGAGACAC ACCTCATCAA
10551 ATGGACTGGG AACGTGGGAA GGGACAGGAC CTGATGTCCC CTTTACTCTC
10601 CCCTCTTCTG GCTCTGCGT TCCCTCTGCG TGCCCCAGTC TGTGTTCCGG
10651 AACTTGACG TCGATGGGAA TGCCACATC TCACAGGAAG AATTCCAGAT
10701 CATCCGTGGG AACTCCCTT ACCTCAGCGC CTTTGGGAC CTCGACCAGA
10751 ACCAGTGGG AGGGCTGGGG ACCTGGGGAA GAGGGAAAGGC AACTCAGCCC
10801 ACTTCTGCCT GGGCTTCAGT TTCTTGTGTG CAAGATGAGG TCACTGAGCC
10851 AGATGATCTT GGCCTGGGAA GCTGCCAGTG TGGGAAAGGG CACTTGCTTT
10901 TGTGGGGAGG AGAGGCTGCC AGCTGTGGAG GCGCAGTGGT ATCTCACAAA
10951 TTCAGACAGA TGGGGGGCTC CACCTGAGTC TTGCAAAGAC TGTGACCTGG
11001 GGACTGTGGC TACAAAAGTG CTGTTTATT TGTGGAGCTC ACAGCTGTCA
11051 AGAAGTGTGG GCAACTTGAG CTCCTGGATA GTCTGTTCTA ATGAATAGAT
11101 AAGAAAGGTT TGTAATTAGC AGTACCCAGT TGTATTCAA CAGTTCATAT
11151 GCTGACAATT TGGAAAACA GCTGGTCTC TGAAGTAGGT TAAACATGCC
11201 CCCTGAAGCC AGATTCTAGC CCTATTTCCTG CTGAGCAGAA AAAACTCCAT
11251 TCAAAATTAA AAGTCCATCT CAGGTCGATT TATTTTTAA TGTTACCTGT
11301 ATTCAAAAAA TCTGTTGTTT TTTATTCCA CATTACAAA ATCCACGGTA
11351 AAATAAAATC TAGTGGTAA AATAAATTAA TAGTGAACAA AATGTTAAA
11401 GTAAGAAGTG AGAGGCCAGG TGCGGTGCCT CACGCCCTGTA ATCCTAGCAC
11451 TTTGGGAGAC TGAGTTGGCA GGATCAATTG AGGCCAGGAG TTTGAGCCCA
11501 GCCTGGCAA CAGAGTAAGA CCCTGTCTCT ACAAAAATTA TTATTATTAT
11551 TTTTGAGACA GAGTCTCACT CTGTTGCCCT GGCTGGAGTG CAGTGGTACA
11601 ATCTCGGCTC GCTGCAACCT CCACCTCCTG GTTCAAGTG ATTCTCCCTG
11651 TTCAGCTTCC TGAGTAGCTG GGATTACAGG CATGCATCAC CGTGCCTGGC
11701 TAATTTTGT ATTTTAGCA GAGATGGGGT TTTACCATGT TGGCCAGGCT
11751 GGTCTCAAAC TCTTGACCTC AAGTGTACTA CCTGCCTTGG CCCCCCAAAG
11801 TGCTAGGATT ACAGGCATGA GCTACTGCTC CTAGCCTAAA AAAATTTTT
11851 TTGGGCATGG GTGGCACGTG CCTGTAGTCC CAGCTACTCA GGAGGCTGAG
11901 GCAGGAGGAA CCCTTGAGCC CAGGAGGTTG AGACTGCAGT GAGCTGTCA
11951 CACACCACTG CACTTCAGGC TGGGTGACTG CGCGAGATCA CCCCCATCAA
12001 AAAAAAAA AAAAGAAAAA AAAAGGAAGA AATGAAAGTC CCCTCTTCC
12051 TTTTCCACTG GTAGAAGTT CCATGATTA GCACTGTTAA CAATATTAAG
12101 CTTGGCAGTA TGTGGATTCT TCCAGTCTTC TTTTCCCAAGG CAGGTGCACA
12151 TTGATAGAGA TTTTGTGTTGT TTGGTGTCTG TTTCATGGAC AAACAGGATT
12201 AGAGCATAAA TCTAGTTCTG CTTGTGGCTT TTATCATAGC TGCTTTATTT
12251 CTTCTCCCAG ATTTTAGGCA GAGGTAGTTG AGTTCCATGT TTTCTCCCTG
12301 GGTTGGTGGG TGGATTTTA TCTAGACCAC CTTTCAGTG AGAATGACCC
12351 TTTGAGACGA TGGAGGCCTC AGCTTCATGC AGCGGGCTCA GCCTTAACCC
12401 TCCACCTCCT GCAGGGCCCCA AGCTGTGTGT GTGTGTGTGT GTGTGTGT
12451 GTGTGTGTGT GTGTGTGTGT GTTGGTAAGG GGAAAGCCCC TGGTTGGGTA
12501 TCAAAACCT AGCACCTGGT TCGGCAGGAG GGAGACCAGC ACCGGCTCCC
12551 CAGGACCAGG CCCAGCTCAC CACTTCATTG TAAAGCTCCC TCTTTGTTTC

FIGURE 3, page 4 of 12

12601 TGGAACCTGG GTGTTCCAT TTCTTCTTA CAAAATTATC TATGCATTTA
 12651 CAGCAATTGT TGATATATCT TTAGGCAGCA TCTAGGTACT TGTAGTGGGT
 12701 TCTCTTTTT CTTTTTTCTT TTTTTTAATC ACCCTCTT TTTTTTGAGA
 12751 CAGAGTCTCA CTCTGTCGCT CAGGCTGGAG TGAATAGCG CGATCTTGGC
 12801 TCACTGCAAC CTCTGCCTCC CAGGTTCAAG TAATTCTCAT GCCTCAGCCT
 12851 CCCAAGTAGC TGAGATTACA GGCACGGCC ACCAGACCCG GCTAATT
 12901 TTTCTTTTT CTTTTTTTG AGACGGAGTT TCGCTCTTG TTGCCCAGGC
 12951 TGGAGTACAG TGGTGTGATC TCGGCTCACT GCAACCTCCG CCTCCCGGGT
 13001 TCAAGTGTATT CTCCTGTCAG AGCCTCCGA GTAGCTGGGA TTACAGGCGC
 13051 GCGCCACCAT GCCTGGCTAA TTTGTATT TTTTTTTTTT GAGACAGAGT
 13101 CTCACTCTGT CACCCAGACT GGAGTGCCTT GGCACGATCT CGGCTCACTG
 13151 CAAGCTCTGC TTCCCCGGTT CATGCCATT CTCGCCTCA GCCTCCGGAG
 13201 TAGCTGGGAC TACAAGCACC CACCAACCGT CCCGGCTAAT TTTTGATATT
 13251 TTTAGTAGAG ACGGGGTTTC ACCGTGGTCT CGACCTCCAG ACCTCGTGT
 13301 CCACTAGCCT CAGCCTCCCCA AAGTGCCTGG ATTACAGGCG TGAGCCACCT
 13351 CACCCAGCCT AATTTGTAT TTTTAGTACA GATGGGGTTT CACCATGTTG
 13401 CGCAGGCTGG TATTGAACCT CTGACCTCA GTGATCCGCC CGCCTCGGCC
 13451 TCCCCAAGTT CTGGGATTAT AGGCGTGAGC CACCGCACCT GGCCTAATT
 13501 TTGTATTTT AGTAGAGATG GAGTTTACC TTGTTGGCCA GGCTGGCTT
 13551 GAACTCCTGA CCTCACCTCA GGTGATCTGC CCACCTCGGC CTCCCAAAGT
 13601 GCTGGGATTA CAGGCATGAG CCACTGTCA CCCGGCCTAA AAATCACCCT
 13651 CTTGACAGAA CTTCACGCC TGTCTTTGT TTTTTTTCAT CTTTGTGCTT
 13701 GTTTCACACT TAACCCCTGA TCACAGACAT CTTCCATGT GGATTCACTG
 13751 AGAACTACCT CATTGTTAG AACAGCTGCA GAGTATTCCA CTGTGCGTT
 13801 AGTCATCAT TTCCCTAACCT ATCCTCCTGC TGATGGACAG TTAGACTGTT
 13851 CCAGTTTTTC AGTATGATTCT TATGCCAGGC TGCCATGAAC GTCCTTTAC
 13901 TGATCCACTC AGGCCAGTAT TTCTGTAGGA GAAATTCTA GAAAGTGGGAT
 13951 AATTGGATCA AAAGATATGC ACATTCTAAA TTAGGAGAGA GACTGCCAAA
 14001 CTGACCTCAAG ACCAGTTGTG ACCAGTTGC ACCCCCCATCA GCAGCGTACA
 14051 AGTGCCTGCT TCCCAACTC CTCGCCAACCA GGGATGCTAT AAAAAGCTTC
 14101 ACAATTTCAGC CAGTCTCAT GGCAAAATGGT ATCTTGGTTA AATTGCAATT
 14151 TCTTTAACAC TAAGTGGGGG TAGGGTATCT TTTCATATGT TTATTGGCCA
 14201 TTATTTCTT CTGTCAATTG CCTGTTCTGA TTCCCTGTCC ATTATTCTAC
 14251 TGGGTTGTT GGTCTTTTC TCATTGATTT TTAGAATCTC TGTAAATGGA
 14301 TATTAACCCCT TTGCTGTGAA ATGTGTTGC AAATATTTC TCCCTGTCTG
 14351 TCATTTATGT GTCTTTTCC ATATAAAATT AAAAATTTT GGTGGGCTCA
 14401 ATAGGTCAGT CTTTCCCTTC CGGGCTTCTG GGATTGTGT TCAGGGTAGA
 14451 AAGGCCCTCA GCCCCCTCAAG ATTATAAAAT TATAAAACCT TTTCTTTT
 14501 TTTTTTTTCTGAGACAGG GTGTCTTGCC ATGTCACCCA GGCTGGAGTG
 14551 CAGTGGCATG ATCTGGCTC GCTGCAACCT CCACCTCCCA GGTCAAAGTG
 14601 ATTCTCGTGC CTTAGCTCC CGAGTAGCTG GGATTATAGG TGCCTGCCAC
 14651 TATGCCCTGGC TAATTCTTGT TATTTTTAGT AGAGACGGGG CTTTGCCATG
 14701 TTGGCCAGGC TGCTGTGAA CTCCGTACCT CGTGATCCAC CGCCTTGCG
 14751 CTCCCAAAGT GCTGGGACTA CAGGCGTAAG CCACTGTGCT CGGCCCTATA
 14801 TTTTTTCAG ATAGCCAGTT ATCCTAATGC TCCCTTGATT TGATGGACCA
 14851 CCTGGATCAC ACATTATGAG CCCCCCTCATA AGCAGGTGGG AGTCTCAAGC
 14901 GAGGCCAGT CCCGATGGGA ATAGCACTTG GTGGCTGAGG ACCCTCCAT
 14951 CTGTGCAGAC ACTGTGTAA AACTTCACAT GCATCATCTA ATTTAGTCT
 15001 CACCAAAATC CTATGAAATG TAGGAATGAT CATTACACCC ATTTATAGAT
 15051 AAGGAAACGG AGGGACAGGG AGATTACTCC GCTACAGGTC AAGAGGCAGG
 15101 GAAGTAGAGC TGCGATTTGA ACTGAGGTCT GTGTCTAGAA CACGTGCTCA
 15151 TTCTTCCCT AAAATGTATT CATAGGTGAA AAAGGGCTTC TGCGGAAAGC
 15201 CCTGGGTTAT GTGGGAAACCT CGGATTTCAG AGCTGTCTT CCAGCAGGAT
 15251 GATGCAGGAG AGAGAGGGAT GCGATTCTC CCAATCTCTC CTGGTCCCAG
 15301 AACTCATTAG AGAGTTCTCC CTGCTGAGGG CTCCCGACTG GTGTTGCACA
 15351 CAGTACACTT CGGGAGCCCG AGGCTGATGG TTCCATGGAA AGTACACAGT
 15401 CATTAGTT TGACACCAA GTGTGAAGTG GGCAGGACAG GCCACTGTT
 15451 TGAGAAGGAA CCCAGGGAAA GGGACTGGCC CAAGACCAACA CACTGGTAG
 15501 CGGCACTTCC CACATCTGCC TGACCCCTAG TCCAGTGCCG CCTTTCTT
 15551 ACTCTGCAAC AGGAGTCAA AATCAGGAGT TCCATGAGGA CACTGGGAAC
 15601 AGTGGGATGG GTTACGGCCAG CGGTGGATGG TTCTGGGGAG GGCCCGAGCT
 15651 GAAGCGCCCC CGCAACTCCC CACAGGGATG GCTGCATCAG CAGGGAGGAG
 15701 ATGGTTTCCT ATTTCCTGCG CTCCAGCTCT GTGTTGGGGG GGCGCATGGG

15751 CTTCGTACAC AACTTCCAGG AGAGCAACTC CTTGCGCCCC GTCGCCCTGCC
15801 GCCACTGCAA AGCCCTGGT AGAGTCCCTT TCCCGGCTCA CGGGCCAAGC
15851 CACGCCCTC CAGCCCCGGC CCCGCCCTCC CTCTGGCCC CGCCCTCTGCC
15901 AGAGCCCTTC TCAAGCCAGG AAAACCTGGT AATTCTATTT GCCTCTCCTC
15951 CTGTGGTTCT GCCCGGGGGC CTGAGGGGG CTCTAAAGCC CTAGTCTCAC
16001 CCTCAAGAAG GAAGAAGTAG AGTCATCACC TCTAAATCCC TCCTCCCACC
16051 ACGGCCCTC CTCTATTGCA GATCCTGGC ATCTACAAGC AGGGCCTCAA
16101 ATGCCGAGGT GAGATGGAAT GACTGGAAGG GCTGCTGGC AGTGTTTTT
16151 TTGTTTGTGTT GTTGTTGG GAGAGTTACT ATTTTGGTGG GGCAATTGCC
16201 AAGGAGTGAA GTACCTTAAA ATCAGAGGCG CATGGCCGGG CATGGTGGCT
16251 CAAGCCTGTA ATCCCAGCAC TTTGGGAGGC CGAGGCGCGC AGATCACCTG
16301 AGGTCAAGGAG TTCAAGACCA GCCTGACCAA CATAGCGCAA CCCCGCCTCT
16351 ACTAAAAATA CAAAAAGTAG CTGGGCGTGG TGCGACCCAC CTGTAATCCC
16401 AGCTACTTGG GAGGCTGAGG CATGAGAATC GCTTGAACCT GGGAGGCGGG
16451 GTTTGCAGTG AGCCGAGATC ACGCCACTGC ACTCCAGCCT GGGCAACAGA
16501 GAGGGCTCTG TCTCAAAAAA AAAAAACAAAC AAAAAAAACCC CAAAACCAA
16551 AACCCCACAA AATCAGAGGC TCAAGATGAC TGATGTGAAG GGAGTGGCGT
16601 TTAAGAGGCC ATTTATTTTG ATGACGCAGC TGCCCAGGAA CAGAGAACAT
16651 GGGAGAAGGC ATAGACTGAC AATTAGGAGG AGGAGAACAC TTTGGAAGGA
16701 GACTCTTATT TTGGTGGGGC AGCTGCTAG GAACAAAGGT TCCTGGTAGG
16751 GGGGCGCAAG CCTGCGGGGAT GGGATGGAGG GTATTCTGAC CAATGTCCCT
16801 GGCTGGCTCT CCATTTGCTC TCCCCCAGCC TGTGGAGTGA ACTGCCACAA
16851 GCAGTCAAG GATCGCCTGT CAGTTGAGTG TCGGCGCAGG GCCCAGAGTG
16901 TGAGGCTGGA GGGGTCTGCA CCCTCACCCCT CACCCATGCA CAGCCACCAT
16951 CACCGCGCCT TCAGCTTCTC TCTGCCCGC CCTGGCAGGC GAGGCTCCAG
17001 GCCTCCAGGT AAGAGGGAGT CATTCTGTAC TGCCCTGTGG AGGGAAAGGAT
17051 GCAGGGCTAC TGGGGCAAAG AACGCAGGAT GGAAGCCATT CAAAAGTGCA
17101 TAATTCTCTT TTGTTGGTGG GATAATAAAG AAGGGACAGG CCGGGCGCGG
17151 TGGCTCACGC CTGTAATCCC AGCACTTTGG GAGGCGCGAGG CGGGCGGATC
17201 ACGAGGTCAG GAGATCGAGA CCATCTTGGC TAACACGGTG AAACCCCATC
17251 TTTACTAAAA ATACAAAAAA AAAAAAATTAG CCAGGCGTGG TGGCGGACGC
17301 CTGTAGTCCC AGCTACTTGG GAGGCTGAGG CAGGAGAATG GCATGAACCC
17351 GGGAGGCGGG GCTTGCAGTG AGCCGAGATC GCGCCACTGC ACTCCAGCCT
17401 GGGCGATAGA GCAAGACTCC GTCTCAGAAA AAAAAAAAAT AAAAAATAAA
17451 GAAGGGACAG GTAAGGGTGC CAGAAAGTGG CCAGGAAGCC CTGGACCTTC
17501 TGAGGCTGAG GAGAGAGACC CTAATTATA AAGAGGTATA AAAGTGAAG
17551 AGGCTTCAAG ATTCCAGTTA CAGTCTTATT TTGTTGGAGG GGTTAACAAA
17601 GGATTGGAGA AGGTGTTATA TGAGCCATTG GCTTGCCTTT CCCTTCTGG
17651 CTGCTCTGGA GGCTCTTCTG GGGAAAGTCC CTTGCCCTGA TAATGTCTG
17701 GCAGCTCTCT TGGGGTATTG GATGGTTTA GGTCAAGTTG CTGAATGACA
17751 ACTGCCAAA TGATTATTT GCTGAGAACAA GTCCGAACAA CTATGTTAAA
17801 CTGGGGTCTA AGGTAGTTGA TCACAACGTG TTGGGTTGGC ATAAGTCTC
17851 AAAAAACAGA GGCAGGCACA GGGCATAACAT CCTCAAAAT AGAAAAGATA
17901 AATCATTG CATTGAGCCT TCCAGAAGTG CTGGGGTCTA AAATGTGAAA
17951 TACACACAAA ATTGACATTG AAGCAAACCT CGCTGACAAA TCTGTGGCTG
18001 AAAAGCTGT GGCAAAACAA AAACATAGAA AAAGAGCCTC AAAAATTGGG
18051 CTGAGGCGAG GCATGGTGGC TCACGCTGT AATCTAGCA CTTTGGGAAG
18101 CCAAGGTGGG TGGATCACCC GAGGTCAAGA GTTGGAGACC AGACTGGCCA
18151 ACCTGGCAAA ACATCATCTC TACAATACAA AAATACAAA ATTAGCTGGG
18201 CGTGGTGGCA GGCGCCTGTA ATCCCAGCTA CTTGGGAGGC TGAGGCACGA
18251 GAATCGCTTG AACCTGGGAG GTGGAGGTTG CAGAGAGCCG AGATTGCGCC
18301 ATTGCACTCC AGCCTGGGGC ACAGAGAGAG ACTCTGTCTC AAAAAAAA
18351 AAAAAAAA AAATTGGGCT GTGAGGTCA GCAGGGAATT GATTTTGGT
18401 GGGTGGGTCT GCTTCTGGGA TGATGTGGAT GCCTCCCGTG GAGAGGGAA
18451 GGGTTGATGA AGTCCCAGGG ACCTGGAAGT GTGTTCTGCA GCAATCCCC
18501 TCCCAGCAGA GATCCGTGAG GAGGAGGTAC AGACGGTGGA GGATGGGTG
18551 TTTGACATCC ACTTGTAAATA GATGGTGAGT CCTCCCACAG CTGGCACCAG
18601 AGCTCCCCAC TGAGGGCTGG GGGGGAGCTG GGGAGTATCA GGGAAATGGG
18651 TGCTTATCC AAATGGCTCC AAGCCAGGTG GGCTACTACC TTGTTGTAG
18701 GGGGGTGTCT TCCTCACAACT CTGTTTTCT CTTCCCAGCT GTGGTTGGAT
18751 CAAGGACTCA TTCTGCCTT GGAGAAAATA CTTCAACCAG AGCAGGGAGC
18801 CTGGGGGTGT CGGGGCAGGA GGCTGGGAG GGGGGTGGGAT TATGAGGGTG
18851 GCATGCAGCT GAGGGCAGGG CCAGGGCTGG TGTCCTAAG GTTGTACAGA

18901 CTCTTGTGAA TATTGTATT TTCCAGATGG AATAAAAAGG CCCGTGTAAT
 18951 TAACCTTCAC CATCAGCGCC TAGAACATCCCG GGGGGTAGGG GGATGCTATA
 19001 CTTTACAGGA TGACAATCTT GGGAGCTAGA ACTTTGTAGC CAGAGAAACT
 19051 TGGGAGGTCT GGAATCTCAT GTGTCTGGAG TCTTGGGGAA GAGAATCTTA
 19101 GAAGCAGAAA ACCTTGGAAC ATAAGAATCT TGGGGAGGGT CTAGGATCTT
 19151 GAGGGAGACCA GATCCTTGGG CATCTAAAC TTGAAACTAG TAGGTCTGCA
 19201 CCCGAGAAATT GCAGGGCCAG TCATGCATAC CCAAAGCCTT CAGCCCATGG
 19251 CCGAAATTCC CTTGCTGGAC AGGGGGCCTT TCAGCCCCCTG CTTGGACGCT
 19301 TCCAGTAACA GGGCCCTCAC TGCAAGAAC GTGGGAGGGG GAGGGGCAGC
 19351 ACAGAGTTGC TGGCTGTCGG GGAAGGGAGG GAGGGCCCTG GGCAGTCGA
 19401 GGGCCCTGCT GGGCTTGTGC CTCAGGGTGG GGGCTGCACT CCTCCGCCTT
 19451 GCAGCCTCCT GGCCTGGTGC TGCTGCCAGC CGGAAGGACA GTGACTTCCA
 19501 GAGGAAATGC ATATTGATCC TGCTTCAGC CTCCGGTGGT GGCTTCTCCC
 19551 AACCCAGCTC TTCCCTCCTG AGCCTGCAGC ACGGAGGTTT TGGGGTCAC
 19601 TGCTACCTAA AGAAGGCTAA GGCCACTTCT GAGGGCTGGC TGGGAGTTTA
 19651 CTAAAGGTTTC TGAAGCTGGG CGGGGCTGCC CCTGGGATCA GGAGACTCCA
 19701 GACAGCAGTC CTGACAATGG GAACTACCTC CTCAGTCCCC CAAACTGGG
 19751 GGTGTCCCAC AGCAGCTGTA GGATTGTCCT AGGGGTGGAG ACCTGAGCAC
 19801 CTTCCACTCC AAAGCACAGT ATCTGTGGG CTGGCAGTGG CCTCAGTTCC
 19851 CCCATGAGTG CCCCCGGTCCC CCACCCAGG GTTCCCCCAC ATCACATCCA
 19901 TCCCTGCTTT GAGACCCAC TCCCCCTGGC CTGTTCTTTA TTTTGGGTCA
 19951 CTCCCTTCTC TTTCCTGGTC ATATCTCTCC TGCAAGGCCA CCCTGTGTTG
 20001 GGCCCCCAG CCCTGTCTCT GCATGGGTG CCCCCCTGCC CCTCCCTCTG
 20051 TCCTCAGCCCC CCTCCGCCCT TCCCCCTCTT GAGGCTGTAA TATCCGTTTC
 20101 ACGATTGAGGG GGCTGAGTTG CTATAACAAAC AGACGGCGAT TGTGTTGTGA
 20151 AGAGCAGCTC GCTCCTGTGC CGCCTGCCCTC CTGTCGTGCC TCCATCCCTG
 20201 CAGCCCAGTC GGTTCCCTCTT GGCTCCTCTC GTCACTACCC TCCAGTTCCA
 20251 GTCTGGCCTC TTCTGGTGT GTGTGTTGT GTGTGTTGTAT GTGTGTTGT
 20301 GCATGCATGC ATATGTTGTGT CCAGGTCTGC CTGCCCCGGGA TGTGACAAGT
 20351 AGCGGTCTTC ATGGTTGCAT GTGTCTGAAT TTGGTGTCTG AGCTTCACAT
 20401 TGTATGCCCTC TGTGTGCATG TGTGTGCATG GACATGCATG CTGTATCTGC
 20451 TGTGTTTCCC CTCCCCCATG TGTCCCCACT GGCTTTGCA CATGGGAGAA
 20501 GGGCATGTGC TCAGCATATC ACTCAACTGT CCACATTGGG TGGGTACCTG
 20551 TGTGTTGTGT GTGTGTTGTGG GGGGTGTGTC TTGAAGTGGC AGGTCCAAA
 20601 TGCTTAGGCA ATCTGAACCT TGGACCTTGC AGAGAGGAGA GATGTCCCTG
 20651 TAGGTGGGAG GGACAGGGAG ATGCAGCAGC TGCCCCGGTGA CCTTTCTGC
 20701 CCTTGATGGG CAAAGCTGGG GGTAGGGAAA GGAGACAAAGT GCTCATACTT
 20751 ACCTCCCTCC CTGCCCCAGGC TCCTCTGTAA GGGTCTGAGT CTGTCTCTGT
 20801 GAGCCATTGC ATCTGTCTGT CTATGCCCTG ATGCCCTGGAT GGACAAGGGG
 20851 TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG AGTGTGAGGC TGCAGGAAGA
 20901 GGAACAGTGG GGGATGGGCA GGAAAGTGGG CTGTGGGGTC AGGGAGGCAGA
 20951 T (SEQ ID NO:3)

FEATURES:

Start: 3000
 Exon: 3000-3072
 Intron: 3073-3753
 Exon: 3754-3855
 Intron: 3856-4363
 Exon: 4364-4427
 Intron: 4428-4786
 Exon: 4787-4918
 Intron: 4919-5702
 Exon: 5703-5853
 Intron: 5854-6056
 Exon: 6057-6230
 Intron: 6231-6389
 Exon: 6390-6506
 Intron: 6507-8832
 Exon: 8833-9114
 Intron: 9115-9885
 Exon: 9886-9963

Intron: 9964-10201
 Exon: 10202-10324
 Intron: 10325-10638
 Exon: 10639-10754
 Intron: 10755-15675
 Exon: 15676-15817
 Intron: 15818-16071
 Exon: 16072-16108
 Intron: 16109-16828
 Exon: 16829-17008
 Intron: 17009-18491
 Exon: 18492-18565
 Stop: 18566

CHROMOSOME MAP POSITION:

Chromosome 11

ALLELIC VARIANTS (SNPs):

DNA

| Position | Major | Minor | Domain |
|----------|-------|-------|----------------|
| 5539 | C | G | Intron |
| 5658 | T | G | Intron |
| 5861 | C | T | Intron |
| 6023 | A | G | Intron |
| 6799 | C | T | Intron |
| 9579 | C | A | Intron |
| 9842 | T | C | Intron |
| 10159 | T | C | Intron |
| 12025 | A | - G | Intron |
| 14723 | T | C | Intron |
| 14996 | G | A | Intron |
| 16153 | T | G | Intron |
| 16181 | G | A | Intron |
| 16756 | A | G | Intron |
| 18059 | A | G | Intron |
| 18364 | A | - T | Intron |
| 18861 | G | A | Beyond ORF(3') |
| 20443 | G | A | Beyond ORF(3') |
| 20881 | A | T | Beyond ORF(3') |

Context:

DNA

Position

5539 AGACTCAGTCAGGCATGAAGTCTCCGTGGGCTTGAGGGTTGGGGCTTCCGGGTA
 GAATTTGTCGTTCCCACCTCTGTTTCCATGGCACTTTGTACAGACTCCTGTACAAAGAC
 CTCTGTACATGTGTCACGCTGTTGTGATCATGTGTTCTGTCTGTCTCCCTCAGTA
 GACTGTGAGCTCTCGAGGGCAGGAACCGTGTCTACTCATCTGTATTCCCAGGCC
 AGCACAGTGCCTGGCACAGAGTACGTTGTTCATAAATGTGTGTTGAGTGCATGACGGGT
 [C, G]
 GGGGGAGATGAGGGAGGAGTTGCTGGACTGGAACATTCTGCCTAGGACAGTGCCTCGC
 ATTATGTAGGTTCTCAGTAAGCGTGAATGGTGTGTTGAGTGGGGGCCACGAGGC
 ATGCGCATGTCAGCAAGGGCTACTACCCCTGCCCTACCTACAAGTGG
 AGCGGCAGGTGACTCAGCGGAACCCGTGGGACAGAAAAAGCGCAAGATGTCCCTGTTGT
 TTGACCACCTGGAGCCATGGAGCTGGCGGAGCATCTCACCTACTTGAGTATCGCTCCT

 5658 CCTCTGTACATGTGTCACGCTGTTTGTGATCATGTGTTCTGTGTCTGTCTCCCTCAGT
 AGACTGTGAGCTCCTCGAGGGCAGGAACCGTGTCTACTCATCTGTATTCCCAGGCC
 TAGCACAGTGCCTGGCACAGAGTACGTTGTTCATAAATGTGTGTTGAGTGCATGACGGGG
 TGGGGGGAGATGAGGGAGGAGTTGCTGGGACTGGAACATTCTGCCTAGGACAGTGCCTC
 GCATTATGTAGGTTCTCAGTAAGCGTGAATGGTGTGTTGAGTGGGGGCCACGAG

[T, G]
 CATGCGCATGTCCAGCAAAGGGCTCACTACCCCTGCCCCCCCAGCCCTACCTACAAGTGG
 AAGCGGCAGGTGACTCAGCGAACCTGTGGGACAGAAAAAGCGCAAGATGTCCCTGTTG
 TTTGACCACCTGGAGCCCATGGAGCTGGCGGAGCATCTCACCTACTTGGAGTATCGCTCC
 TTCTGCAAGATCCTGGTGCGGCCCCAGGGCTGGGGTCAAGGGTCAATGTGGGCTGGA
 AGAGAGTTCTAGGAGGGCAGGGTCCCTGGCTAGGCTGGTACAGGGTGCATCAGGGG

5861 CTGGGACTGGAACATCGTGCCTAGGACAGTGCCTCGCATTATGTAGGTTCTCAGTAAG
 CGTGAATGGTGTCTGAGTGGGGGCCACGAGGCATGCGCATGTCAGCAAAGGG
 CTCACTACCCCTGCCCCCCCAGCCCTACCTACAAGTGGAAAGCGCAGGTGACTCAGCGGA
 ACCCTGTGGACAGAAAAAGCGCAAGATGTCCCTGTTGACCACCTGGAGGCCATGG
 AGCTGGCGGAGCATCTCACCTACTGGAGTATCGCTCCTCTGCAAGATCCTGGTGCAGGC
 [C, T]
 CGAGGGCTGGGGTCAAGGGTCAAATGTGGCTGGAAGAGAGTTCTAGGAGGGCAGGG
 TCCCTGGCGTAGGCTGGGTACAGGGTGATCAGGGTTTCAGTGTAACTGAAGGTC
 AGCTGGAGGGTAGGGTGGCTATCAGTGAGGGAGAGGGCGGCAAGGTGCTGAGGCCAC
 TCCTCATGCCAGTTTCAAGGACTATCACAGTTCTGACTCATGGCTGCAGTGGAC
 AACCCCGTCTGGAGCGGTCATCTCCCTTCAACAGCGTCAACAGTGGTGCAGCTC

6023 GGCAGGTGACTCAGCGAACCTGTGGGACAGAAAAAGCGCAAGATGTCCCTGTTGTTG
 ACCACCTGGAGCCCATGGAGCTGGCGGAGCATCTCACCTACTTGGAGTATCGCTCCTTCT
 GCAAGATCCTGGTGCAGGGCGAGGGCTGGGGTCAAGGGTCAAATGTGGCTGGAAGAG
 AGTTCTAGGAGGGCAGGGTCCCTGGCTAGGCTGGTCAAGGGTGATCAGGGTTTC
 AGTGTAAACCACTGAAGGTAGCTGGAGGGTGAGGAGTGGCTATCAGTGAGGGAGAGGCC
 [A, G]
 GCAAGGGTCTGAGGCCACTCCTCATGCCCCCAGTTCAAGGACTATCACAGTTCTGACT
 CATGGCTGACTGTGGACAACCCCGCTGGAGCGTTCATCTCCCTTCAACAGCGTC
 TCACAGTGGGTGCAGCTCATGATCTCAGCAAACCCACAGCCCCGAGCGGGCCCTGGTC
 ATCACACACTTGTCCACGTGGCGAGGTGCCTGCCCTCCGTGTCTCCAACC
 AACCCACATGCCAGTCAGGCCAACCTCCCTCCCTAACCCACTGCCTCTCTAGA

6799 CCATCAAGGTGCCTGGACTGGGAGGGGCCGGTCTTCCCAGGTCTGTCTTCACTGGGT
 CCTCCCAGCAGCACTGGGGCTGGGCACAGCTGTCTCATTTGATAGATATGAAATGGA
 GGCTCAGAGGGGTTAAGTGTCTTCAAGTTGACAATGGCAACAGCAGAGTGGGGCT
 CACAGGTGTCAGGGACCCCAAAGCTAGTACTTTTTTTTTAAGACAGGGTCTC
 TCTCTGTGTCAGACTGGAGTTCAAGTGGTCAGTCACAAGCTACTGCAGCCTGAA
 [C, T]
 TCCCTGAGCTCAATCGATCCTCCCACCTCAGCCTCTGAGTAGCTGGACTACAGGTGTAC
 GCCACCATGCCCTAATTTGTATTGTTAAATTTTTTTTTAGAGATGGGG
 TTTTGCCTGTTGCCCAGACTGGCTTGAACCTCTGGCTCAAGTGATCCGCCTGCCCTG
 GCCTCCAAAGTGTGAGATTATGGCTGAGCCATTGTGCCACTGTAGTTCTT
 CTTTCTTCTCCTCATTTTATTATTTGAAGTATTGAGTAACATA

9579 CTGGTTGGGGCTGTGGACTGAGGCTGATCTCCAGGCTGGTATGTGGACTGTGGCAGTT
 TGAACCTGGGCTGGGCCCCGGGTTGAGTTCTGGCAATGGGCTGTGTTCTAGGGCTGGGCC
 AAGCTCTGCATTCTGGGGCAGGGGTTCTAAGCATGGCCCTGGGCTGGAGTGAAG
 TTCTGGGCTTGGCTTACACTTGGCTTGGGCTAGGGTGGAGTTGGTTCTGGTTA
 GATCCAGACAAGGTTCTAGACATTGGCTGGGCTTAAGTGTAAAGTTGGAGTGGATT
 [C, A]
 TTAGCTGCTTCTGGGCTCTGGAGGGGATCAGGGTTGAAATCAGAGCTCTGGCTGGGTT
 CGACCTGGCTTCTCCCTGACATCTGGCAATATGTTGTTCAAGGTTGGGCCATGC
 TGTGGTTGATCTGTGGCTGGGATGACATGGGGTTGCTGTGCTGTTCTAAGCCAGG
 CTTTGTCTGAGTCTAGCTCTGACCCGAGCTCTGGCTGAGCTGTGCCCTAGGTGAC
 CTTTGGCCCTGGGCTGTGGCCGTGGCAGGGCCAGTGGGGTGTACAGATCTGTGTG

9842 TGGGCTGGGCTTAAGTGTAAAGTTGGAGTGGATTCTAGCTGCTCTGGCTCTGGA
 GGGGATCAGGGTTGAAATCAGAGCTCTGGCTGGGTCGGACCTGGCTTCTCCCTGACA
 TCTTGGCAATATGTTGTTCAAGGTTGGGCCATGCTGTGTTGATCTGTGGCTGG
 GATGACATGGGGTTGCTGTGCTGTTCTAAGCCAGGCTTGTGCTGAGTCTAGCT
 GACCCGAGCTCTGGCTGAGCTGTGGCCTCTAGGTGACCTTGGCCCTGGGCTGTGG
 [T, C]
 GTGGGCAGGGGCCAGTGGGGTGTACAGATCTGTGTCAGGTTGCTCTGGATCAGTA
 TCAGACGGAGGATGAGCTGTACCTGAGCTGTCAGCAGTGGCTGCCAGCAGCTCCAAGTCCTC

GGTGAGGGGGTACTCCCTCCCTCCACTCTGCCCTTCCCTCCTGAGAATCCAGGATGTG
 AGGATGGGAAGAGCTCTAGCAGCACCTACCCATCCATCTTGTAGGACAGAGGCATCC
 TGGGGTAGGGCAGTAGTGTGGCAGACTTCCCTCTCCAGGGATTCCCTCTCTGTTC
 10159 GGGGGTAGATCTGTGTCCCAGGTGTCCTGGATCAGTATCAGACGGAGGATGAG
 CTGTACCACTGTCCTGCAGCGGAGCGCCTCAAGCTCTCGGTAGGGGGTACTCC
 CTCCCTCCACTCTGCCCTCCCTCCTGAGAATCCAGGATGTGAGGATGGAAAGAGCTC
 TTAGCAGCACCTCACCCATCCATCTTGTAGGACAGAGGCATCTGGGGTAGGGCAGTA
 GTGTTGGCAGACTTCCCTCTCCAGGGATTCCCTCTCTGTTCCTGGCTCTGGCT
 [T, C]
 CCCCTGCCCTGGCCCTAGCTCAGGCCGACCACTTCATGCCAACAGCCCCACGAGT
 TGCAACCCACCACCCGGCCCCGGTACTGGAGGAGTGGACCTCGGTGCCAACCCAAG
 CTGGATCAGGCCCTGTGGTGGAGCACATCGAGAAGATGGTGGAGGTGAGCTCCTGCGGA
 GCCTGAGCAGTGTGTGGGAGAGGCCAGTTGCCGGAGCACTGCCCTGGAAGCCAGCACG
 AGTGTCCCTGTCAGAACGCCAGCACTCAGCCCCCTAGGAGTCACAGGGCTGGCAGGCCAGC
 12025 TGGGGTTTACCATGTTGCCAGGCTGGTCTAAACTCTTGACCTCAAGTGTACCTACCTG
 CCTTGGCCCCCAAAGTGTAGGATTACAGGCATGAGCTACTGCTCTAGCCTAAAAAAA
 TTTTTTTGGCATGGTGGCACGTGCCGTAGTCCCAGCTACTCAGGAGGCTGAGGCAG
 GAGGAACCTTGAGGCCAGGAGGTTGAGACTGCAGTGTGAGCTACACCACTGCCT
 TCAGCCTGGTGACTGCCGAGATCACCCATCAAAAAAAAAAGAAAAAAA
 [A, -, G]
 GAAGAAATGAAAGTCCCTCTTCCACTGGTAGAAGTTGCCATGATTAAGCACT
 GTTAACAATATTAAGCTTGGCAGTATGTGGATTCTTCCAGTCTTCTTCCCAGGCAGGT
 GCACATTGTAGAGATTGTTGTTGGTCTGTTGATGGACAAACAGGATTAGAGC
 ATAAATCTAGTCTGTTGTGGCTTATCATAGCTGCTTATTCTCTCCAGATT
 AGGCAGAGGTAGTTGAGTTCCATGTTCTCCCTGGTTGGTGGGATTTATCTAG
 14723 GGCTTCTGGGATTGTGTTGGGGTAGAAAGGCCCTCAGCCCCCTCAAGATTATAAAA
 TAAAACCTTTCTTTTTTTCTGAGACAGGGTGTCTGCCATGTCACCCAGG
 CTGGAGTGCACTGGCATGATCTGGCTCGCTCACCTCACCTCCAGGTCAAGTGT
 TCTCGTGCCTAGCCTCCGAGTAGCTGGATTATAGGTGCCTGCCACTATGCCCTGGCTA
 ATTTTTGTATTTAGTAGAGACGGGCTTGCATGTTGCCAGGCTGGTCTGA
 [T, C]
 CTGACCTCGTGTACCAACCCGCCCTGGCCTCCAAAGTGTGGACTACAGGCGTAAGCCA
 CTGTGCTGGCCCTATATTTTCAGATAGCCAGTTATCCTAATGCTCCCTGATTGA
 TGGACCACCTGGATCACACATTATGAGCCCTCATAACAGCAGGTGGGAGTCTCAAGCGAG
 GGCCAGTCCCAGGGAAATAGCACTGGTGGCTGAGGACCTCCTATCTGTGAGACACT
 GTTGTAAAACCTCACATGCATCATCTAATTAGTCCTACCAAAATCTATGAAATGTAG
 14996 CCATGTTGCCAGGCTGGCTTGAACCTCTGACCTCGTGTACCCACCCGCTTGGCCTCCC
 AAAGTGTGGGACTACAGCGTAAGCCACTGTGCTGCCCTATATTTTCAGATAGC
 CAGTTATCTAATGCTCCCTGATTGATGGACCCACTGGATCACACATTATGAGCCCC
 TCATAAGCAGGTGGAGTCTCAAGCGAGGGCAGTCCCGATGGGAATAGCACTTGGTGGC
 TGAGGACCTCCTATCTGTGCAAGACACTGGTAAAACCTACATGCATCATCTAATT
 [G, A]
 TCCTCACCAAAATCTATGAAATGAGGAATGATCATTACACCCATTATAGATAAGGAA
 ACGGAGGGACAGGGAGATTACTCCGCTACAGGTCAAGAGGCAGGGAAAGTAGAGCTGGC
 TTGAACGTAGGTCTGTCTAGAACACAGTGTCTATTCTTCCCTAAATGTATTCTAGG
 TGAAAAGGGCTCTGGGAAAGCCCTGGGTTATGTGGGAAACCTGGATTACAGCTGT
 CTTTCAGCAGGATGATGCAGGAGAGAGAGGGATGCGATTCTCCAAATCTCCTGGTC
 16153 CGCCCTCCAGCCCCGGCCCGCCCTCCCTCTGGCCCCGCTCTGCCAGAGCCCTTCTC
 AAGCCAGGAAACCTGTAATTCTATTGCCCTCCTCTGTGGTCTGCCGGGGCCCT
 GAGGCGGGCTCTAAAGCCCTAGTCTCACCCCTCAAGAAGGAAGTAGAGTCATCACCTC
 TAAATCCCTCCCTCCACCACGCCCTCCTCTATTGCAAGATCCTGGCATCTACAAGCAG
 GGCCCTCAAATGCCAGGTGAGATGGAATGACTGGAAGGGCTGCTGGCAGTGT
 [T, G]
 TTTGTTGTTGGAGAGTTACTATTTGGTGGGCAATTGCCAAGGAGTGAAGTA
 CCTTAAATCAGAGGCCATGCCGGCATGGTGGCTCAAGCCTGTAATCCAGCACTTT
 GGGAGGCCAGGCCAGATCACCTGAGGTCAAGGAGTTCAAGACCAGCCTGACCAACAT
 AGCGCAACCCGCCCTCACTAAAAACACAAAAGTAGCTGGCAGTGGCACCACCTG
 TAATCCAGCTACTTGGGAGGCTGAGGATGAGAATCGCTTGAACCTGGGAGGCCGGGTT

16181

CTTCTGGCCCCGCCTGCCAGAGCCCTTCTCAAGCCAGGAAAACCTGGTAATTCTATTT
GCCTCTCCCTCTGTGGTTCTGCCGGGCCCTGAGGCGGGCTCTAAAGCCCTAGCTCAC
CCTCAAGAAGGAAGAAGTAGAGTCATCACCTCTAAATCCCTCCACGGCCCTC
CTCTATTGCAGATCCTGGGCATCTACAAGCAGGGCCTCAATGCCAGGTGAGATGGAAT
GACTGGAAGGGCTGCTGGCAGTGTGTTGTTGTTGGAGAGTTACT
[G, A]
TTTGGTGGGCAATTGCCAAGGAGTGAAGTACCTTAAATCAGAGGCCATGGCCGGC
ATGGTGGCTCAAGCCTGTAATCCAGCACTTGGGAGGCCAGGCCAGATCACCTGA
GGTCAGGAGTTCAAGACCCAGCCTGACCAACATAGCGCAACCCGCCCTACTAAAATAC
AAAAAGTAGCTGGGCGTGGCACCCACCTGAATCCAGCTACTTGGGAGGCTGAGGC
ATGAGAATCGCTGAACCTGGGAGGCCGGTTGCAGTGAGCCGAGATCACGCCACTGCA

16756

CAGTGAGCCGAGATCACGCCACTGCACTCCAGCCCTGGCAACAGAGAGGGCTCTGCTCA
AAAAAAAAAAACAACAAAAACCCCCAAAACCAAAACCCACAAAATCAGAGGCTCAAG
ATGACTGTGTAAGGGAGTGGCTTAAAGAGGCCATTATTTGATGACGCACTGCCC
AGGAACAGAGAACATGGGAGAAGGCATAGACTGACAATTAGGAGGAGGAACACTTGG
AAGGAGACTCTTATTTGCTGGGCAGTCAGGAACAAAGGTTCTGGTAGGGGGC
[A, G]
CAAGCCTGCCGGATGGATGGAGGGTATTCTGACCAATGCCCTGGCTGGCTCTCCATT
GCTCTCCCCCAGCCTGTGGAGTGAACGCCACAAGCAGTGCAGGATGCCCTGTCAGTT
AGTGTGGCGCAGGGCCCAGAGTGTGAGGCCCTGGAGGGCTGCACCCCTCACCCA
TGCACAGCACCACATCACCGCCCTCAGCTCTCTGCCCGCCCTGGCAGGCCAGGCT
CCAGGCCTCCAGGTAAGAGGGAGTCATTCTGACTGGCCTGTGGAGGAAGGATGCAGGG

18059

AATGATTATTTGCTGAGAACAGTCGAACAACATGTTAAACTGGGGCTAAGGTAGTT
GATCACAACCTGTTGGGTGGCATTAAGTCCTAAAAAACAGAGGCCAGGCACAGGGCATA
ATCCTCAAAAATAGAAAAGATAAAATCCATTGCAATTGAGCCTTCCAGAAAGTGTGG
TAAATGTGAAATACACACAAAATTGACATTAAAGCAAACCTGCGCTGACAAATCTGTGG
TGAAAAGCTGTGGAAAACAAAACATAGAAAAGAGCCTCAAAATTGGGCTGAGGCC
[A, G]
GGCATGGTGGCTCACGCCGTGAACTTGGCAAGGCAAGGTGGGTGGATCACC
CGAGGTAGGAGTTGGAGACCAGACTGCCAACGTGGAAAACATCATCTACAATACA
AAAATACAAAATTAGCTGGCGTGGTGGCAGGCCCTGTAATCCAGCTACTTGGGAGG
CTGAGGCACGAGAACGCTTGAACCTGGGAGGTGGAGGTGAGAGAGGCCAGATTGCG
CATTGCACTCCAGCCTGGGCACAGAGAGAGACTCTGCTCAAAAAAAAAAAAAAA

18364

TGGTGGCTCACGCCGTGAACTTGGCAAGGCAAGGTGGGTGGATCACCGAG
GTCAGGAGTTGGAGACCAGACTGCCAACGTGGAAAACATCATCTACAATACAAA
TACAAAATTAGCTGGCGTGGTGGCAGGCCCTGTAATCCAGCTACTTGGGAGGCTGA
GGCACGAGAACGCTTGAACCTGGGAGGTGGAGGTGAGAGAGGCCAGATTGCCATT
GCACTCCAGCCTGGGCACAGAGAGAGACTCTGCTCAAAAAAAAAAAAAAA
[A, -, T]
TGGGCTGTGAGGTGAGGTCATGCAGGGAAATTGATTTGGTGGGTGGATGATG
GTGGATGCTCCCGTGGAGAGGGGAAGGGTTGATGAAAGTCCCAGGGACCTGGAAAGTGT
TCTGCAGCAATCCCCCTCCAGCAGAGATCCGTGAGGAGGAGGTACAGACGGTGGAGG
GGGGTGTGACATCCACTTGTAAATAGATGGTGAATGCTCCTCCACAGCTGGCACCAGAGCT
CCCCACTGAGGGCTGGGGAGCTGGGAGTATCAGGGAAATGGGTGCTTATCAAAT

18861

ACTTGTAATAGATGGTGAAGTCCCTCCACAGCTGGCACAGAGCTCCCACTGAGGGCTGG
GGGGGAGCTGGGAGTATCAGGGAAATGGGTGCTTATCCTAAAGGCTCCAAGCCAGGTG
GGCTACTACCTTGTGTTAGGGGGTGTCTCCTCACAAACCTGTTTCTCTCCAGCT
GTGGTTGGATCAAGGACTCATTCTGCCCTGGAGAAAATACTTCAACCAGAGCAGGGAGC
CTGGGGGTGCGGGCAGGAGGCTGGGAGTGGGATATGAGGGTGGCATGCAGCT
[G, A]
AGGGCAGGGCCAGGGCTGGTGTCCCTAACGGTTGTACAGAGACTCTGTGAATATTTGATTT
TCCAGATGGAATAAAAGGCCGTGTAATTAAACCTTCACCATCAGGCCCTAGAATCCCG
GGGGTAGGGGGATGGTATACTTACAGGATGACAATCTGGGAGCTAGAACTTTGTAGCC
AGAGAAACTTGGGAGGTCTGGAATCTCATGTGCTGGAGTCTTGGGAAAGAGAAATCTTAG
AAGCAGAAAACCTTGGAACATAAGAATCTGGGAGGTCTAGGATCTGAGGAGACCAG

20443

TGTTGTGAAGAGCAGCTCGCTCCTGTGCCGCCCTGCCCTGTGCTGCCCTCCATCCCTGCA
GCCCAAGTCGGTCTCTTGGCTCCTCGTCACTACCCCTCAGTCCAGTCTGGCCTCTT

CCTGGTGTGTGTGTGTGTGTGTGTGTGTGTATGCATGCATGCATATGTGTGTCC
AGGTCTGCCCTGCCCGGATGTGACAAGTAGCGGTCTTCATGGTTGCATGTGTCTGAATTT
GGTGTCTGAGCTTCACATTGTATGCGCCTGTGTGCATGTGTGCATGGACATGCATGCT
[G, A]
TATCTGCTGTGTTCCCTCCCCATGTGTCCCCACTGGCCTTGCACATGGGAGAAGGG
CATGTGCTCAGCATATCACTCAACTGTCCACATTGGGTGGGTACCTGTGTGGTGTGTG
TGTGTGGGGGTGTCTTGAAGTGGCAGGTCCAAATGCTTAGGCAATCTGAACCTGG
ACCTTGCAGAGAGGAGAGATGCCCTGTAGGTGGGAGGGACAGGGAGATGCAGCAGCTGC
CCGGTGACCTTTCTGCCCTGTAGGGCAAAGCTGGGGTAGGGAAAGGAGACAAGTGCT
20881 TTGAAGTGGCAGGTCCAAATGCTTAGGCAATCTGAACCTGGACCTTGCAGAGAGGGAGA
GATGTCCCTGTAGGTGGGAGGGACAGGGAGATGCAGCAGCTGCCGGTGACCTTCTGC
CCTTGATGGCAAAGCTGGGGTAGGGAAAGGAGACAAGTGCTCATACTTACCTCCCTCC
CTGCCAGGCTCCTCTGAAGGGCTGAGCTGTCTCTGTGAGCCATTGCATCTGTCTGT
CTATGCCCTGATGCCCTGGATGGACAAGGGTAGTGTGTGTGTGTGTGTGTGTGTG
[A, T]
GTGTGAGGCTGCAGGAAGAGGAACAGTGGGGATGGCAGGAAAGTGGCTGTGGGTCA
GGGAGGCAT